

EQUIVALENCE CLASSES AND LOCAL ASYMPTOTIC NORMALITY IN SYSTEM IDENTIFICATION FOR QUANTUM MARKOV CHAINS

MADALIN GUTA AND JUKKA KIUKAS

ABSTRACT. We consider the problems of identifying and estimating dynamical parameters of an ergodic quantum Markov chain, when only the stationary output is accessible for measurements. On the identifiability question, we show that the knowledge of the output state completely fixes the dynamics up to a ‘coordinate transformation’ consisting of a multiplication by a phase and a unitary conjugation of the Kraus operators. When the dynamics depends on an unknown parameter, we show that the latter can be estimated at the ‘standard’ rate $n^{-1/2}$, and give an explicit expression of the (asymptotic) quantum Fisher information of the output, which is proportional to the Markov variance of a certain ‘generator’. More generally, we show that the output is locally asymptotically normal, i.e. it can be approximated by a simple quantum Gaussian model consisting of a coherent state whose mean is related to the unknown parameter. As a consistency check we prove that a parameter related to the ‘coordinate transformation’ unitaries, has zero quantum Fisher information.

1. INTRODUCTION

Quantum system identification has recently received significant attention due to its relevance in understanding complex quantum dynamical systems and the development of quantum technologies [13]. Among the different setups under investigation, we mention channel tomography [16], Hamiltonian identification [6, 11], and the estimation of the Lindblad generator of an open dynamical system [27].

In this paper we study two system identification problems set in a context which is particularly relevant for quantum control engineering applications [29, 34], namely the input-output formalism of quantum Markov dynamics, which in the continuous time framework has been studied in [17]. Our setup is that of a discrete time quantum Markov chain consisting of a quantum system (or ‘memory’) $\mathcal{H} = \mathbb{C}^D$ interacting successively with input ancillas (or ‘noise units’) $\mathcal{K} = \mathbb{C}^k$ which are identically prepared in a known pure state $|\chi\rangle \in \mathcal{K}$ (see Figure 1). The goal is to learn about the dynamics encoded in the unitary operator $U : \mathcal{H} \otimes \mathcal{K} \rightarrow \mathcal{H} \otimes \mathcal{K}$, by measuring the quantum output consisting of the noise units after the interaction. We will assume that we do not have direct access to the ‘memory’ system, and we do not control the input state, both assumptions corresponding to realistic experimental constraints, e.g. in atom maser experiments [23]. We will also assume that the system’s transition operator is irreducible and aperiodic, so that in the long time limit the dynamics has forgotten the initial state and has reached stationarity.

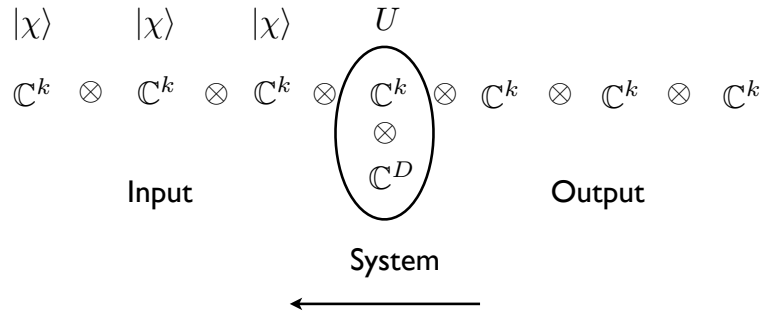


FIGURE 1. Illustration of a quantum Markov chain. A sequence of identically prepared input ancillas (or ‘noise units’) with space \mathbb{C}^k interact successively (from right to left) with a system (or ‘memory’) \mathbb{C}^D via the unitary U . After n interactions the output ‘noise units’ are correlated with each other and with the system, and carry information about the unitary U .

In the stationary regime, the output is in a finitely correlated state [15] denoted ρ_V^{out} which is completely determined by the isometry $V : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{K}$ mapping the system into a system-ancilla space in one evolution step

$$V : |\varphi\rangle \mapsto U(|\varphi\rangle \otimes |\chi\rangle).$$

In particular, the restriction of the output state to the first n noise units in the stationary regime is given by $\rho_V^{out}(n) := \text{tr}_{\mathcal{H}}[V(n)\rho_{ss}V(n)^*]$, where ρ_{ss} is the system's stationary state and $V(n) : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{K}^{\otimes n}$ is the n -steps iteration of the isometry V (see below for the precise definition).

The first problem we address is that of finding the *equivalence classes* of isometries with identical output states. In Theorem 2 we show that $\rho_{V_1}^{out} = \rho_{V_2}^{out}$ if and only if V_1 and V_2 are related by an arbitrary complex phase c and a conjugation with a local unitary W on the system \mathcal{H}

$$V_2 = c(W^* \otimes \mathbb{1}_{\mathcal{K}})V_1W.$$

We point out that a similar result holds in the following ‘classical’ setup. A (finite) hidden Markov chain is a discrete time stochastic process $(X_n, Y_n)_{n \geq 0}$ consisting of an ‘underlying’ Markov chain $(X_n)_{n \geq 0}$, with state space $\mathcal{X} := \{1, \dots, D\}$ and transition matrix $[T_{i,j}]_{i,j=1}^D$, and a sequence of random variables $(Y_n)_{n \geq 0}$ with values in $\mathcal{Y} := \{1, \dots, k\}$, which ‘depend’ on the underlying Markov dynamics. More precisely, $(Y_n)_{n \geq 0}$ are independent when conditioned on $(X_n)_{n \geq 0}$, and Y_n depends on the Markov chain only through X_n , with conditional distribution $R_{a,i} = \mathbb{P}(Y_n = a | X_n = i)$. We consider that $(X_n)_{n \geq 0}$ is ‘hidden’ (not accessible to observations) and we observe the sequence $(Y_n)_{n \geq 0}$ whose marginal distribution depends only on the matrices T and R and the initial state π of the Markov chain

$$\mathbb{P}(Y_1 = a_1, \dots, Y_n = a_n) = \sum_{i_0, \dots, i_n} \pi(i_0) \prod_{j=0}^{n-1} T_{i_j, i_{j+1}} \prod_{j=0}^n R_{i_j, a_j}.$$

Under ergodicity conditions similar to ours, and certain additional generic conditions, Petrie [30] has shown that by observing $(Y_n)_{n \geq 0}$ in the stationary regime, we can identify the matrices (T, R) up to a permutation of the labels of the hidden states in \mathcal{X} . Since the output process of a quantum Markov chain can be interpreted as a quantum analogue of a hidden Markov process, our result can be seen as a quantum extension of [30].

After tackling the identifiability question, the second problem we consider is that of estimating the isometry V , or more precisely finding how the output state $\rho_V^{out}(n)$ varies with V , and how much statistical information it contains. Our approach is based on asymptotic statistics, and in particular on the concept of local asymptotic normality [33]. To illustrate this in the case of the hidden Markov model, we assume that the matrices (T, R) depend smoothly on an unknown real parameter θ which we would like to estimate, and denote by \mathbb{P}_{θ}^n the probability distribution of the data (Y_1, \dots, Y_n) . For large n the unknown parameter can be ‘localised’ in a neighbourhood of size $n^{-1/2}$ of a fixed value θ_0 (e.g. a rough estimator), such that $\theta = \theta_0 + u/\sqrt{n}$ with u a ‘local parameter’ to be estimated. Local asymptotic normality means that the statistical model $\mathbb{P}_{\theta_0 + u/\sqrt{n}}^n$ can be approximated by a simple Gaussian model consisting of a single sample from the normal distribution $N(u, I_{\theta_0}^{-1})$ with mean u and fixed variance [4]. Closely related to this, and perhaps more relevant for the practitioner, is the fact that the maximum likelihood estimator $\hat{\theta}_n = \hat{\theta}_n(Y_1, \dots, Y_n)$ is asymptotically normal [12, 5], i.e.

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} N(0, I_{\theta_0}^{-1})$$

where the convergence holds in distribution, when $n \rightarrow \infty$. In particular, the mean square error $\mathbb{E}[(\hat{\theta}_n - \theta)^2]$ scales as $1/n$ with the best possible constant $I_{\theta_0}^{-1}$ which is the inverse limiting Fisher information per sample of the hidden Markov chain, at θ_0 .

Returning to our quantum system identification problem, we note that in order to find good estimators we need to deal with the optimisation problem of finding the ‘most informative’ output measurement. In quantum state tomography this problem has been approached by developing quantum analogues of several key statistical notions such as quantum Fisher information [25, 24], and quantum local asymptotic normality [28, 19]. Equipped with these mathematical tools it is possible to solve the optimal state estimation problem in an asymptotic regime, by transforming it into a simpler Gaussian estimation one [18]. We will therefore apply the same strategy for the Markov system identification problem, but we note that since the output ‘noise units’ are not independent, we cannot use the existing state tomography theory in a straightforward fashion. Instead, we developed a notion of convergence of *pure states* statistical models based on the simple idea that two models are close to each other if the inner products of all corresponding pairs of vectors are very similar. To illustrate this, let $|\Psi_{u,\varphi}(n)\rangle := V_{\theta_0 + u/\sqrt{n}}(n)|\varphi\rangle$ be the joint system and output state after n steps, with initial condition $|\varphi\rangle \in \mathcal{H}$. In Theorem 4 we show that

$$\lim_{n \rightarrow \infty} \langle \exp(-iau^2)\Psi_{u,\varphi}(n) | \exp(-iav^2)\Psi_{v,\varphi}(n) \rangle = \langle \sqrt{F/2}u | \sqrt{F/2}v \rangle = \exp(-F(u-v)^2/8)$$

where the right side is the inner product of two coherent states of the limit Gaussian model $|\sqrt{F/2}v\rangle$, with quantum Fisher information F . The latter can be computed explicitly and can be interpreted as a certain Markov variance of the ‘generator’ $G = -iV\dot{V}^*$, similarly to the well known formula for unitary rotation

families of states. Building on this result, we show that the convergence to the Gaussian model can be extended to the mixed stationary output state itself, and can be formulated in a strong, operational way by means of quantum channels connecting the models in both directions [28]. In this statistical picture, the local unitary conjugation corresponding to changes inside an equivalence class has quantum Fisher information equal to zero, in agreement with the fact that such parameters are not identifiable.

The paper is organised as follows. In section 2 we review the formalism of quantum Markov chains, discuss the ergodicity assumptions and introduce the ‘Markov covariance’ inner product, which will be used later in interpreting the quantum Fisher information. In section 3 we prove Theorem 2 which characterises the equivalence classes of chains with identical outputs. Section 4 is a brief introduction to the statistical background of the paper, centred around the notion of convergence of quantum statistical models, and local asymptotic normality. The main result here is Lemma 5 which can be used to convert weak convergence of pure states models into strong convergence. Section 5 contains several local asymptotic normality results for different versions of the output state. In all cases, the quantum model described by the output state can be approximated by a quantum Gaussian model consisting of a one parameter family of coherent states. Some of the more technical proofs are collected in section 6.

A special case of the present local asymptotic normality result has been obtained in [22]. A detailed study comparing the quantum Fisher information with classical Fisher informations of various counting statistics for the atom maser model can be found in [9, 8]. The continuous time version of the local asymptotic normality result will appear in a forthcoming publication [7].

2. QUANTUM MARKOV CHAINS

We begin by describing the general framework quantum Markov chains, and establishing some of the notations used in the paper. In the second part of this section we introduce a positive inner product describing the variance of Markov fluctuation operators, which will be used later for interpreting the limiting quantum Fisher information.

2.1. Output state, Markov transition operator and the quantum Perron-Frobenius Theorem. A discrete time quantum Markov chain consists of a ‘system’ with Hilbert space $\mathcal{H} := \mathbb{C}^D$ which interacts successively with ‘noise units’ (or ancillas, quantum coins) with identical Hilbert spaces $\mathcal{K} := \mathbb{C}^k$, cf. Figure 1. The noise units are initially prepared independently, in the same state $|\chi\rangle \in \mathcal{K}$, and the system has initial state $|\varphi\rangle \in \mathcal{H}$. The interaction is described by a unitary operator $U : \mathcal{H} \otimes \mathcal{K} \rightarrow \mathcal{H} \otimes \mathcal{K}$, such that after one step, the joint state of system and the first unit is $U(|\varphi\rangle \otimes |\chi\rangle)$, while the remaining units are still in the initial state. After the interaction with the first noise unit, the system moves one step to the *left* and the same operation is repeated between system and the second unit, and so on. Alternatively, one can think that the system is fixed and the chain is shifted to the right. After n steps the state of the system and the n noise units is therefore

$$|\Psi(n)\rangle := U(n)|\varphi \otimes \chi^{\otimes n}\rangle \in \mathcal{H} \otimes \mathcal{K}^{\otimes n}$$

where $U(n)$ is the product of unitaries $U^{(n)} \dots U^{(1)}$ with $U^{(i)}$ denoting the copy of U which acts on system and the i -th noise unit, counting from *right to left* according to the dynamics of Figure 1.

Our goal is to investigate how the output state (reduced state of the noise units after the interaction) depends on the unitary U and in particular which dynamical parameters can be identified by performing measurements on the output. With this in mind we note that the state $|\Psi(n)\rangle$ can be expressed as

$$|\Psi(n)\rangle = V(n)|\varphi\rangle$$

where $V(1) \equiv V : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{K}$ is the isometry given by $V|\varphi\rangle := U|\varphi \otimes \chi\rangle$, and $V(n) : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{K}^{\otimes n}$ is obtained iteratively from $V(k+1) := (V \otimes \mathbb{1}_{\mathcal{K}^{\otimes k}})V(k)$. Therefore, since the input state is fixed and known, the output state depends on U only through the isometry V which will be the focus of our attention for most of the paper.

Let us fix orthonormal bases $\{|e_j\rangle\}_{j=1}^D$ and $\{|i\rangle\}_{i=1}^k$ in \mathcal{H} and respectively \mathcal{K} , and let $\mathbf{K} := \{K_i\}_{i=1}^k$ be the collection of Kraus operators acting on \mathcal{H} , uniquely defined by the equation

$$V|\varphi\rangle = \sum_{i=1}^k K_i |\varphi\rangle \otimes |i\rangle. \quad (1)$$

which satisfy the normalisation condition

$$\sum_{i=1}^k K_i^* K_i = \mathbb{1} \quad (2)$$

Conversely, any set $\mathbf{K} := \{K_i\}_{i=1}^k$ satisfying (2) has a unique associated isometry $V : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{K}$ given by (1). The system-output state can be written more explicitly in a matrix product form

$$|\Psi(n)\rangle = \sum_{\mathbf{i} \in I^{(n)}} \mathbf{K}_{\mathbf{i}}^{(n)} |\varphi\rangle \otimes |\mathbf{i}\rangle,$$

where $I^{(n)}$ denotes the set of multi-indices $\mathbf{i} = (i_1, \dots, i_n)$, and for each $\mathbf{i} \in I^{(n)}$ we denote by $\mathbf{K}_{\mathbf{i}}^{(n)}$ the n -step Kraus operator (note the backwards ordering)

$$\mathbf{K}_{\mathbf{i}}^{(n)} := K_{i_n} K_{i_{n-1}} \cdots K_{i_2} K_{i_1}.$$

For later purposes we define the concatenation of multi-indices $\mathbf{j} \in I^{(k)}$ and $\mathbf{i} \in I^{(m)}$ via

$$\mathbf{ji} := (i_1, \dots, i_m, j_1, \dots, j_k) \in I^{(k+m)},$$

so that $\mathbf{K}_{\mathbf{ji}}^{(k+m)} = \mathbf{K}_{\mathbf{j}}^{(k)} \mathbf{K}_{\mathbf{i}}^{(m)}$.

The properties of the system-output state $|\Psi(n)\rangle$ depend crucially on the irreducibility of the restricted system dynamics. Let $\mathcal{A} := \mathcal{B}(\mathcal{H})$ and $\mathcal{B} := \mathcal{B}(\mathcal{K})$ be the algebras of system and noise observables, and let us denote by $\mathcal{A}_*, \mathcal{B}_*$ their preduals (linear spans of density matrices). In the Schrödinger picture the reduced evolution of the system is obtained by iterating the transition operator (quantum channel)

$$\begin{aligned} T_* &: \mathcal{A}_* \rightarrow \mathcal{A}_* \\ T_* &: \rho \mapsto V \rho V^* = \sum_{i=1}^k K_i \rho K_i^* \end{aligned}$$

as it can be seen from the identity

$$V(n) \rho V(n)^* = \sum_{\mathbf{i} \in I^{(n)}} \mathbf{K}_{\mathbf{i}}^{(n)} \rho \mathbf{K}_{\mathbf{i}}^{(n)*} = T^n(\rho).$$

The dual (Heisenberg) evolution is given by the unit preserving CP map

$$\begin{aligned} T &: \mathcal{A} \rightarrow \mathcal{A} \\ T &: X \mapsto V^* X \otimes \mathbb{1}_{\mathcal{K}} V = \sum_{i=1}^k K_i^* X K_i. \end{aligned} \quad (3)$$

As for classical Markov chains, the system's space may possess non-trivial invariant subspaces, and have multiple stationary states. We will restrict our attention to such 'building blocks' defined by the following properties.

Definition 1. A channel $T : \mathcal{A} \rightarrow \mathcal{A}$ is called

- (i) *irreducible* if there exists $n \in \mathbb{N}$ such that $(\text{Id} + T)^n$ is strictly positive, i.e. $(\text{Id} + T)^n(X) > 0$ for all positive operators X .
- (ii) *primitive*, if there exists an $n \in \mathbb{N}$, such that T^n is strictly positive.

We call an isometry $V : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{K}$ irreducible (primitive), if the associated channel defined in (3) is irreducible (primitive).

Clearly primitivity is a stronger requirement than irreducibility. The following theorem (see [14, 32]) collects the essential properties of irreducible (primitive) quantum transition operators needed in this paper.

Theorem 1 (quantum Perron-Frobenius). *Let $T : \mathcal{A} \rightarrow \mathcal{A}$ be a completely positive map and let $r := \max_i |\lambda_i|$ be its spectral radius, where $\{\lambda_1, \dots, \lambda_{D^2}\}$ are the (complex) eigenvalues of T arranged in decreasing order of magnitude. Then*

- (i) *r is an eigenvalue of T , and it has a positive eigenvector.*
- (ii) *If additionally, T is unit preserving then $r = 1$ with eigenvector $\mathbb{1}$.*
- (iii) *If additionally, T is irreducible then r is a non-degenerate eigenvalue for T and T_* , and both corresponding eigenvectors are strictly positive.*
- (iv) *If additionally, T is primitive, then $|\lambda_i| < r$ for all other eigenvalues than r .*

We mainly need the following corollary of this theorem: If a channel T is irreducible then it has a unique full rank stationary state, and if T is also primitive then any state converges to the stationary state ρ_{ss} in the long run (mixing or ergodicity property);

$$\lim_{n \rightarrow \infty} (T_*)^n(\rho) = \rho_{ss},$$

or in the Heisenberg picture

$$\lim_{n \rightarrow \infty} T^n(X) = \text{tr}[X\rho_{ss}]\mathbb{1}. \quad (4)$$

Moreover, the speed of convergence is exponential, the rate depending on the second largest eigenvalue of T .

2.2. The Markov covariance inner product. In this section we introduce an inner product playing the role of a ‘Markov covariance’, whose relevance will become apparent when computing the quantum Fisher information of the output. On a deeper level, we conjecture that the covariance is accompanied by a Central Limit Theorem, but this topic will not be pursued here (see [22, 21] for the special case of output observables).

In this section we assume that V is a primitive isometry, and we denote by $\langle Z \rangle_{ss}^{in}$ the expectation with respect to the input state $\rho_{ss} \otimes |\chi\rangle\langle\chi|^{\otimes n}$. Let $\mathcal{A} \otimes \mathcal{B}^{(j)} \subset \mathcal{A} \otimes \mathcal{B}^{\otimes n}$ be the subalgebra of observables of the system together with the j -th (from right to left) copy of the noise unit \mathcal{B} , and for any $X \in \mathcal{A} \otimes \mathcal{B}$ we denote by $X^{(j)}$ its version in $\mathcal{A} \otimes \mathcal{B}^{(j)}$. If $X \in \mathcal{A} \otimes \mathcal{B}$ we define the Heisenberg evolved operator after i steps by

$$X(i) = U^{(1)*} \dots U^{(i)*} X^{(i)} U^{(i)} \dots U^{(1)} \in \mathcal{A} \otimes \mathcal{B}^{\otimes n},$$

and note that its stationary mean does not depend on i and is given by

$$\langle X(i) \rangle_{ss}^{in} = \langle \mathcal{E}(X) \rangle_{ss}^{in} = \text{Tr}(\rho_{ss} \mathcal{E}(X))$$

where \mathcal{E} is the conditional expectation

$$\begin{aligned} \mathcal{E} : \mathcal{A} \otimes \mathcal{B} &\rightarrow \mathcal{A} \\ X &\mapsto \langle \chi | U^* X U | \chi \rangle = V^* X V. \end{aligned}$$

For all mean zero operators $X \in \mathcal{A} \otimes \mathcal{B}$ we define the associated *fluctuations operator* by

$$\mathbb{F}_n(X) = \frac{1}{\sqrt{n}} \sum_{i=1}^n X(i) \in \mathcal{A} \otimes \mathcal{B}^{\otimes n}.$$

Finally, let $\mathcal{A}_0 := \{A \in \mathcal{A} : \text{Tr}[\rho_{ss} A] = 0\}$ and define $\mathcal{R} : \mathcal{A}_0 \rightarrow \mathcal{A}_0$ to be the inverse of the restriction of the map $\text{Id} - T$ to \mathcal{A}_0 . Note that \mathcal{R} is well defined, as it follows from the fact that T is primitive and has a unique eigenvector $\mathbb{1}$ with eigenvalue 1.

Lemma 1. *Let \mathcal{F} be the linear space of mean zero operators $\{X \in \mathcal{A} \otimes \mathcal{B} : \text{Tr}(\rho_{ss} \mathcal{E}(X)) = 0\}$. For any $X, Y \in \mathcal{F}$ and any $|\varphi\rangle \in \mathcal{K}$ the following limit exists*

$$(X, Y)_V := \lim_{n \rightarrow \infty} \frac{1}{n} \langle \varphi \otimes \chi^{\otimes n} | \mathbb{F}_n(X^*) \mathbb{F}_n(Y) | \varphi \otimes \chi^{\otimes n} \rangle$$

and defines a positive inner product on \mathcal{F} .

Moreover $(X, Y)_V$ has the explicit expression

$$(X, Y)_V = \text{Tr} \{ \rho_{ss} \mathcal{E} [X^* Y + X^* (\mathcal{R} \circ \mathcal{E}(Y) \otimes \mathbb{1}) + (\mathcal{R} \circ \mathcal{E}(X^*) \otimes \mathbb{1}) Y] \}.$$

Proof. For simplicity we denote by $\langle \cdot \rangle_\varphi^{in}$ the expectation with respect to the state $|\varphi \otimes \chi^{\otimes n}\rangle$. By expanding the fluctuation operators we obtain

$$\begin{aligned}
& \frac{1}{n} \langle \varphi \otimes \chi^{\otimes n} | \mathbb{F}_n(X^*) \mathbb{F}_n(Y) | \varphi \otimes \chi^{\otimes n} \rangle = \frac{1}{n} \sum_{i,j=1}^n \langle X(i)^* Y(j) \rangle_\varphi^{in} \\
&= \frac{1}{n} \sum_{i=1}^n \langle (X^* Y)(i) \rangle_\varphi^{in} + \frac{1}{n} \sum_{i \neq j=1}^n \langle X(i)^* Y(j) \rangle_\varphi^{in} \\
&= \frac{1}{n} \sum_{i=1}^n \langle \varphi | T^{i-1} \circ \mathcal{E}(X^* Y) | \varphi \rangle + \frac{1}{n} \sum_{1 \leq i < j \leq n} \langle X(i)^* Y(j) \rangle_\varphi^{in} + \frac{1}{n} \sum_{1 \leq j < i \leq n} \langle X(i)^* Y(j) \rangle_\varphi^{in} \\
&= \frac{1}{n} \sum_{i=1}^n \langle \varphi | T^{i-1} \circ \mathcal{E}(X^* Y) | \varphi \rangle + \frac{1}{n} \sum_{1 \leq i < j \leq n} \langle \varphi | T^{i-1} \circ \mathcal{E} [X^* (\mathbb{1} \otimes T^{j-i-1} \circ \mathcal{E}(Y))] | \varphi \rangle \\
&\quad + \frac{1}{n} \sum_{1 \leq j < i \leq n} \langle \varphi | T^{j-1} \circ \mathcal{E} [(\mathbb{1} \otimes T^{i-j-1} \circ \mathcal{E}(X^*)) Y] | \varphi \rangle.
\end{aligned}$$

Now, since $T_*^n(|\varphi\rangle\langle\varphi|)$ converges to ρ_{ss} exponentially fast as $n \rightarrow \infty$, the first term in the last equality converges to $\text{Tr}(\rho_{ss} \mathcal{E}(X^* Y))$. Similarly, the one of the summation indices in the second term can be changed to $k := j - i - 1$ such that it can be written as

$$S_n := \sum_{k=0}^{n-2} \frac{1}{n} \sum_{i=1}^{n-k-1} \langle \varphi | T^{i-1} \circ \mathcal{E} [X^* (\mathbb{1} \otimes T^k \circ \mathcal{E}(Y))] | \varphi \rangle.$$

For any fixed k the inner sum is dominated by terms with large i and by the same stationarity argument as above, it converges to

$$\text{Tr}(\rho_{ss} \mathcal{E} [X^* (\mathbb{1} \otimes T^k \circ \mathcal{E}(Y))]).$$

Now, by assumption, $\mathcal{E}(Y)$ belongs to the domain of \mathcal{R} so that the sum converges

$$\sum_{k \geq 0} T^k \circ \mathcal{E}(Y) = \mathcal{R} \circ \mathcal{E}(Y).$$

The two facts together imply that

$$\lim_{n \rightarrow \infty} S_n = \text{Tr} \{ \rho_{ss} \mathcal{E} [X^* (\mathcal{R} \circ \mathcal{E}(Y) \otimes \mathbb{1})] \}.$$

A similar reasoning applies to the third term of the sum. □

The previous lemma follows from the next conjecture which will not be investigated in this paper.

Conjecture 1 (Central Limit for quantum Markov chains). *Let $X \in \mathcal{A} \otimes \mathcal{B}$ be selfadjoint operator with $\langle X \rangle_{ss}^{out} = 0$. The fluctuations operator $\mathbb{F}_n(X)$ satisfy the Central Limit Theorem*

$$\mathbb{F}_n(X) \xrightarrow{\mathcal{D}} N(0, V_X)$$

where $N(0, V_X)$ is the centred Gaussian distribution with variance $V_X = (X, X)_V$, and the convergence holds as $n \rightarrow \infty$ in distribution with respect to the state $|\varphi \otimes \chi^{\otimes n}\rangle$

The special case where $X \in \mathcal{B}$ has been proven in [22]. Physically, it means that time averages of measurements of the observable X are asymptotically Gaussian.

3. THE EQUIVALENCE CLASS OF QUANTUM MARKOV CHAINS WITH IDENTICAL OUTPUTS

In Theorem 2 of this section we answer the first question set in the introduction: which (mixing) quantum Markov chains have the same output states in the stationary regime?

Definition 2. Let $V_l : \mathcal{H}_l \rightarrow \mathcal{H}_l \otimes \mathcal{K}$, $l = 1, 2$ be two primitive isometries, where $\mathcal{H}_l, \mathcal{K}$ are finite-dimensional Hilbert spaces. Let $\rho_{ss,l}$ denote the respective stationary states. We call V_1 and V_2 *equivalent*, if they have the same output states in the stationary regime, that is,

$$\text{tr}_{\mathcal{H}_1} [V_1(n) \rho_{ss,1} V_1^*(n)] = \text{tr}_{\mathcal{H}_2} [V_2(n) \rho_{ss,2} V_2^*(n)], \quad \text{for all } n \in \mathbb{N}.$$

We begin with a straightforward observation that for any given primitive isometry $V_1 : \mathcal{H}_1 \rightarrow \mathcal{H}_1 \otimes \mathcal{K}$, a number $c \in \mathbb{C}$ with $|c| = 1$, and a unitary $U : \mathcal{H}_2 \rightarrow \mathcal{H}_1$, the isometry $V_2 := c(U^* \otimes I)V_1U$ is equivalent to V_1 , and is primitive (with the stationary state $\rho_{ss,2} = U^*\rho_{ss,1}U$). The following Lemma characterises the case where two given primitive isometries are related this way. The proof is inspired by a similar argument from [1].

Lemma 2. *Let $V_l : \mathcal{H}_l \rightarrow \mathcal{H}_l \otimes \mathcal{K}$, $l = 1, 2$, be two primitive isometries, and define the maps*

$$T_{ll'} : \mathcal{B}(\mathcal{H}_{l'}, \mathcal{H}_l) \rightarrow \mathcal{B}(\mathcal{H}_{l'}, \mathcal{H}_l), \quad T_{ll'}(X) = V_l^*(X \otimes \mathbb{1}_{\mathcal{K}})V_{l'}.$$

for $l, l' = 1, 2$. Then the following conditions are equivalent:

- (i) T_{12} has an eigenvalue of modulus one;
- (ii) T_{21} has an eigenvalue of modulus one;
- (iii) there exists a unitary operator $U : \mathcal{H}_2 \rightarrow \mathcal{H}_1$, and $c \in \mathbb{C}$ with $|c| = 1$, such that

$$V_2 = c(U^* \otimes \mathbb{1}_{\mathcal{K}})V_1U.$$

In that case, $T_{12}(U) = cU$ and $T_{21}(U^*) = \bar{c}U^*$.

Proof. Conditions (i) and (ii) are clearly equivalent: if $T_{ll'}(F) = cF$ with some $c \in \mathbb{C}$ and $F \in \mathcal{B}(\mathcal{H}_{l'}, \mathcal{H}_l)$, then $T_{l'l}(F^*) = \bar{c}F^*$. Assuming (iii) we have

$$T_{12}(U) = V_1^*(U \otimes \mathbb{1}_{\mathcal{K}})V_2 = cV_1^*(U \otimes \mathbb{1}_{\mathcal{K}})(U^* \otimes \mathbb{1}_{\mathcal{K}})V_1U = cU,$$

i.e. (i) holds, with c the corresponding eigenvalue. Thus, the only nontrivial implication is (i) \implies (iii).

Assume (i), let c be an eigenvalue of T_{12} modulus one, and F such that $T_{12}(F) = cF$. Then from the definition of T_{21} it follows immediately that $T_{21}(F^*) = \bar{c}F^*$. Since V_1 is an isometry, we have $V_1V_1^* \leq \mathbb{1}_{\mathcal{H}_1 \otimes \mathcal{K}}$ (in fact, $V_1V_1^*$ is a projection). Hence,

$$\begin{aligned} T_{22}(F^*F) &= V_2^*(F^* \otimes \mathbb{1}_{\mathcal{K}})(F \otimes \mathbb{1}_{\mathcal{K}})V_2 \\ &\geq V_2^*(F^* \otimes \mathbb{1}_{\mathcal{K}})V_1V_1^*(F \otimes \mathbb{1}_{\mathcal{K}})V_2 = T_{21}(F^*)T_{12}(F) = |c|^2 F^*F = F^*F, \end{aligned}$$

so by positivity of T_{22} ,

$$T_{22}^n(F^*F) \geq F^*F \quad \text{for all } n \in \mathbb{N}.$$

Let P be the projection onto the eigenspace of F^*F corresponding to its largest eigenvalue $\|F^*F\|$. Now $\lim_{n \rightarrow \infty} T_{22}^n(X) = \text{tr}[\rho_{ss,2}X]\mathbb{1}_{\mathcal{H}_2}$ by the primitivity of V_2 . Hence,

$$\text{tr}[\rho_{ss,2}F^*F] = \lim_{n \rightarrow \infty} \text{tr}[P]^{-1} \text{tr}[PT_{22}^n(F^*F)] \geq \text{tr}[P]^{-1} \text{tr}[PF^*F] = \|F^*F\|.$$

This implies that $\text{tr}[\rho_{ss,2}F^*F] = \|F^*F\|$, i.e. $\rho_{ss,2}$ is supported in the projection P . But $\rho_{ss,2}$ has full rank in \mathcal{H}_2 , so $P = \mathbb{1}_{\mathcal{H}_2}$, and, consequently, $F^*F = \|F^*F\|\mathbb{1}_{\mathcal{H}_2}$. By proceeding in exactly the same way using the primitive channel T_{11} , we show that $FF^* = \|FF^*\|\mathbb{1}_{\mathcal{H}_1}$. Denote $\alpha := \|FF^*\| = \|F^*F\|$, and $U := \alpha^{-\frac{1}{2}}F$. Then $U : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ is a unitary operator between the two Hilbert spaces and in particular, $\dim \mathcal{H}_1 = \dim \mathcal{H}_2$. Moreover,

$$U^*V_1^*(U \otimes \mathbb{1}_{\mathcal{K}})V_2 = U^*T_{12}(U) = cU^*U = c\mathbb{1}_{\mathcal{H}_2}.$$

For a unit vector $|\varphi\rangle \in \mathcal{H}_2$, this implies

$$\langle V_1U\varphi | (U \otimes \mathbb{1}_{\mathcal{K}})V_2\varphi \rangle = c. \quad (5)$$

But $\|V_1U\varphi\| = \|(U \otimes \mathbb{1}_{\mathcal{K}})V_2\varphi\| = 1$ by isometry, so there is actually *equality* in the Cauchy-Schwartz inequality

$$|\langle V_1U\varphi | (U \otimes \mathbb{1}_{\mathcal{K}})V_2\varphi \rangle| \leq \|V_1U\varphi\| \|(U \otimes \mathbb{1}_{\mathcal{K}})V_2\varphi\|.$$

This is possible only if the two vectors are linearly dependent, i.e. there is a constant $\lambda \in \mathbb{C}$ such that

$$(U \otimes \mathbb{1}_{\mathcal{K}})V_2\varphi = \lambda |V_1U\varphi\rangle.$$

Putting this back in (5), we see that $\lambda = c$. Since φ was arbitrary, we have (ii), and the proof is complete. \square

The following lemma deals with the case where all eigenvalues of the map T_{12} of the preceding lemma have modulus strictly less than one.

Lemma 3. *Let $V_l : \mathcal{H}_l \rightarrow \mathcal{H}_l \otimes \mathcal{K}$, $l = 1, 2$, be two primitive isometries, and define $T_{ll'}$ as in Lemma 2. Let $\rho_l^{\text{out}}(n)$ be the associated output states. Then the limits*

$$\lim_{n \rightarrow \infty} \text{tr}[\rho_1^{\text{out}}(n)^2], \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{tr}[\rho_2^{\text{out}}(n)^2]$$

exist and are strictly positive. If $\lim_{n \rightarrow \infty} T_{12}^n = 0$, or, equivalently, $\lim_{n \rightarrow \infty} T_{21}^n = 0$, then

$$\lim_{n \rightarrow \infty} \text{tr}[\rho_1^{\text{out}}(n)\rho_2^{\text{out}}(n)] = 0. \quad (6)$$

Proof. Let $\rho_{ss,l}$ be the stationary state of T_{ll} , with spectral decomposition

$$\rho_{ss,l} = \sum_{i=1}^{\dim \mathcal{H}_l} \Lambda_{l,i} |e_{l,i}\rangle \langle e_{l,i}|, \quad l = 1, 2.$$

The output states decompose as follows:

$$\begin{aligned} \rho_l^{\text{out}}(n) &= \sum_{i=1}^{\dim \mathcal{H}_l} \Lambda_{l,i} \text{tr}_{\mathcal{H}_l} [|V_l(n)e_{l,i}\rangle \langle V_l(n)e_{l,i}|] \\ &= \sum_{i,j=1}^{\dim \mathcal{H}_l} \Lambda_{l,i} |\psi_{l,ji}(n)\rangle \langle \psi_{l,ji}(n)|, \end{aligned}$$

where

$$\psi_{l,ji}(n) = \sum_{\mathbf{i} \in \mathbf{I}^{(n)}} \langle e_{l,j} | K_{l,\mathbf{i}}^{(n)} e_{l,i} | \mathbf{i} \rangle.$$

Now

$$\begin{aligned} \langle \psi_{l,ji}(n) | \psi_{l',j'i'}(n) \rangle &= \sum_{\mathbf{i} \in \mathbf{I}^{(n)}} \overline{\langle e_{l,j} | \mathbf{K}_{l,\mathbf{i}}^{(n)} e_{l,i} \rangle} \langle e_{l',j'} | \mathbf{K}_{l',\mathbf{i}}^{(n)} e_{l',i'} \rangle \\ &= \langle e_{l,i} | T_{ll'}^n(|e_{l,j}\rangle \langle e_{l',j'}|) | e_{l',i'} \rangle, \end{aligned}$$

so we can write

$$\begin{aligned} \text{tr}[\rho_l^{\text{out}}(n)\rho_{l'}^{\text{out}}(n)] &= \sum_{i,j=1}^{\dim \mathcal{H}_l} \sum_{i',j'=1}^{\dim \mathcal{H}_{l'}} \Lambda_{l,i} \Lambda_{l',i'} |\langle \psi_{l,ji}(n) | \psi_{l',j'i'}(n) \rangle|^2 \\ &= \sum_{i,j=1}^{\dim \mathcal{H}_l} \sum_{i',j'=1}^{\dim \mathcal{H}_{l'}} \Lambda_{l,i} \Lambda_{l',i'} |\langle e_{l,i} | T_{ll'}^n(|e_{l,j}\rangle \langle e_{l',j'}|) | e_{l',i'} \rangle|^2. \end{aligned}$$

Since the isometries V_l are primitive, we have $T_{ll}^n(X) \rightarrow \text{tr}[\rho_{ss,l}X] \mathbb{1}_{\mathcal{H}_l}$ for any $X \in \mathcal{B}(\mathcal{H}_l)$ by (4). Now, on the one hand, by choosing $l = l'$ we get

$$\lim_{n \rightarrow \infty} \text{tr}[\rho_l^{\text{out}}(n)^2] = \sum_{i,i'=1}^{\dim \mathcal{H}_l} \Lambda_{l,i}^2 \Lambda_{l,i'}^2 > 0.$$

On the other hand, assuming $\lim_{n \rightarrow \infty} T_{ll'} = 0$ for $l \neq l'$, we get (6). \square

We are now ready to prove the main result of this section.

Theorem 2. *Two primitive isometries $V_l : \mathcal{H}_l \rightarrow \mathcal{H}_l \otimes \mathcal{K}$, $l = 1, 2$, are equivalent if and only if there exists a unitary operator $U : \mathcal{H}_2 \rightarrow \mathcal{H}_1$, and a complex number c with $|c| = 1$, such that*

$$V_2 = c(U^* \otimes \mathbb{1}_{\mathcal{K}})V_1U. \quad (7)$$

Proof. As mentioned above, the ‘if’ part is straightforward. Assume now that V_1 and V_2 are equivalent, and define $T_{ll'}$ as in Lemma 2. We consider the direct sum isometry

$$V_{\text{tot}} := V_1 \oplus V_2 : \mathcal{H}_1 \oplus \mathcal{H}_2 \rightarrow \mathcal{H}_1 \otimes \mathcal{K} \oplus \mathcal{H}_2 \otimes \mathcal{K} = (\mathcal{H}_1 \oplus \mathcal{H}_2) \otimes \mathcal{K}.$$

We identify the elements $X \in \mathcal{B}(\mathcal{H}_1 \oplus \mathcal{H}_2)$ in the usual way with block matrices

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix},$$

where $X_{ll'} \in \mathcal{B}(\mathcal{H}_{l'}, \mathcal{H}_l)$, the set of linear operators $\mathcal{H}_{l'} \rightarrow \mathcal{H}_l$. This identifies $\mathcal{B}(\mathcal{H}_l, \mathcal{H}_{l'})$ as a subspace $\mathcal{B}(\mathcal{H}_1 \oplus \mathcal{H}_2)$, and each of these four subspaces is invariant under the channel T associated with V_{tot} . Explicitly, we have

$$T(X) = \begin{pmatrix} T_{11}(X_{11}) & T_{12}(X_{12}) \\ T_{21}(X_{21}) & T_{22}(X_{22}) \end{pmatrix}. \quad (8)$$

In particular, any eigenvalue of $T_{ll'}$ is also an eigenvalue of T , because the subspaces are invariant. Since T is completely positive and unital by construction, all eigenvalues of $T_{ll'}$ have modulus at most one. If all eigenvalues of T_{12} have modulus strictly less than one, then $\lim_{n \rightarrow \infty} T_{12}^n = 0$, which according to Lemma 3

contradicts the assumption that the output states are equal. Hence T_{12} has an eigenvalue of modulus one, so Lemma 2 concludes the proof. \square

Remark 1. According to [32], and applying a reconstruction procedure analogous to the one given in [2], one observes that the finitely correlated output states $\rho^{out}(n)$, $n \in \mathbb{N}$, associated to a primitive isometry V are completely determined by one state $\rho^{out}(n_0)$, provided that $n_0 \geq 2(D^2 - d + 1)D^2$, where d is the number of linearly independent Kraus operators. Hence, for any two isometries V_l , $l = 1, 2$, there exists a finite n_0 such that $\rho_1^{out}(n_0) = \rho_2^{out}(n_0)$ if and only if (7) holds for some phase factor c and unitary U .

4. INTERMEZZO ON CONVERGENCE OF QUANTUM STATISTICAL MODELS AND LOCAL ASYMPTOTIC NORMALITY

In this section we introduce the basic elements of a theory of quantum statistical models, in as much as it is necessary to understand the second main result presented in the next section: the local asymptotic normality of the quantum Markov chain's output, and the associated quantum Fisher information. This section is not directly connected to the Markov set-up and can be skipped at a first reading.

Definition 3. Let Θ be a parameter space. A quantum statistical model over Θ is a family

$$\mathcal{Q} := \{\rho_\theta : \theta \in \Theta\},$$

of density matrices ρ_θ on a Hilbert space \mathcal{H} , which are indexed by an unknown parameter $\theta \in \Theta$.

The typical quantum statistical problem associated to a model \mathcal{Q} is to estimate the unknown parameter θ by measuring a system prepared in the state ρ_θ , and constructing an estimator $\hat{\theta}$ based on the measurement outcome. In practice, the problem typically involves an additional parameter n describing the ‘sample size’, and ‘good estimators’ have the property that the estimation error (e.g. the mean square error $\mathbb{E}((\hat{\theta}_n - \theta)^2)$) converges to zero as $n \rightarrow \infty$. The samples may be independent and identical as in quantum state tomography, or may consist of correlated systems depending on an unknown dynamical parameter, as considered in this paper. The *rate* of convergence is typically of the order n^{-1} and has a constant factor equal to the inverse of the (asymptotic) Fisher information, the latter describing the amount of statistical information *per sample*.

Asymptotic statistics deals with the ‘large n ’ statistical inference set-up. The power of this set-up lies in the fact that one can take advantage of general Central Limit behaviour, and approximate the ‘ n samples’ statistical model by simpler Gaussian models, with asymptotically vanishing approximation error [33].

4.1. I.I.D. pure state models. To illustrate this idea, let us consider a simple model consisting of n qubits which are independent and identically prepared in a pure state depending on a two-dimensional rotation parameter

$$|\psi_\theta\rangle := \exp\left(i\sqrt{2}(\theta_2\sigma_x - \theta_1\sigma_y)\right)|0\rangle, \quad \theta = (\theta_1, \theta_2) \in \mathbb{R}^2$$

where the factor $\sqrt{2}$ has been inserted for later convenience. Since we work in an asymptotic framework, we can consider that the parameter θ belongs to a neighbourhood of size $n^{-1/2+\epsilon}$ of a *known* fixed value θ_0 which by symmetry can be chosen to be $\theta_0 = (0, 0)$. Such a ‘localisation’ is not a prior assumption, but can be achieved with an adaptive procedure where a ‘small’ sample $\tilde{n} = n^{1-\epsilon} \ll n$ is used to produce a rough estimate θ_0 , and this information is fed into the design of the second stage optimal measurement [28]. We will therefore write $\theta = u/\sqrt{n} = (u_1/\sqrt{n}, u_2/\sqrt{n})$, where $u = (u_1, u_2)$ is a local parameter to be estimated. In this case, the Gaussian approximation mentioned above is closely related to what is known in physics as the Holstein-Primakov approximation for coherent spin states [26]. By a Central Limit argument one can show that in the limit of large n the collective spin variables

$$L_x := \frac{1}{\sqrt{2n}} \sum_{i=1}^n \sigma_x^{(i)}, \quad L_y := \frac{1}{\sqrt{2n}} \sum_{i=1}^n \sigma_y^{(i)}$$

converge in joint moments (with respect to the product state $|\psi_{u/\sqrt{n}}\rangle^{\otimes n}$) to continuous variables Q and respectively P which satisfy the canonical commutation relations $[Q, P] = i\mathbb{1}$ and have a coherent (Gaussian) state $|u\rangle$ with means $\langle Q \rangle = u_1$ and $\langle P \rangle = u_2$. Therefore, in the large n limit the i.i.d. qubit model

$$\mathcal{Q}_n := \{|\psi_{u/\sqrt{n}}\rangle^{\otimes n} : u \in \mathbb{R}^2\}$$

is approximated (locally around θ_0) by the ‘quantum Gaussian shift’ model

$$\mathcal{G} := \{|u\rangle : u \in \mathbb{R}^2\}.$$

From the statistical viewpoint this approximation (when formulated in an appropriate way) provides the asymptotically optimal measurement procedures and estimation rates [20, 28]. Indeed if we would like to estimate u_1 , then the optimal measurement is that of the total spin L_x which corresponds to the canonical variable Q in the limit model, and similarly for u_2 . However if one is interested in both parameters (i.e. the mean square error is $\mathbb{E}(\|\theta - \hat{\theta}\|^2)$) then the optimal procedure is to measure each L_x and L_y on half of the spins, which corresponds to the heterodyne measurement for the Gaussian model, where the coherent state is split in two, and conjugate canonical variables are measured separately on the two subsystems.

The above convergence to the Gaussian model can be captured in a simple way, as convergence of the inner products for arbitrary pairs of local parameters

$$\lim_{n \rightarrow \infty} \left\langle \psi_{u/\sqrt{n}}^{\otimes n} \middle| \psi_{v/\sqrt{n}}^{\otimes n} \right\rangle = \langle u | v \rangle, \quad u, v \in \mathbb{R}^2.$$

This ‘weak convergence’ [19, 22] has an appealing geometric interpretation and can be verified more easily than the ‘strong convergence’ investigated in [20, 28], the latter being nevertheless more powerful and applicable to more general models of mixed states. At the end of the section we will show how weak convergence can be upgraded to strong convergence under an additional assumption.

4.2. Weak and strong convergence of pure state models. We will now briefly sketch a mathematical framework for weak convergence of pure state models, which will serve as motivation for our results on local asymptotic normality for quantum Markov chains. Since this is not the main focus of the paper, we leave the general theory for a separate work.

Definition 4. Let $\mathcal{Q} := \{\rho_\theta : \theta \in \Theta\}$ be a quantum model with parameter space Θ and Hilbert space \mathcal{H} , and let $\mathcal{Q}' := \{\rho'_\theta : \theta \in \Theta\}$ be another model with the same parameter space and Hilbert space \mathcal{H}' . We say that \mathcal{Q} is equivalent to \mathcal{Q}' if there exists quantum channels T and S such that

$$T(\rho_\theta) = \rho'_\theta, \quad S(\rho'_\theta) = \rho_\theta, \quad \forall \theta \in \Theta.$$

It can be easily seen that if two models are equivalent then for any statistical decision problem, their optimal procedures can be related through the channels T and S and the corresponding risks (figures of merit) are equal [19]. The definition is also naturally connected with the theory of quantum sufficiency developed in [31]. We will now extend this to allow for models which are ‘close’ to each other but not necessarily equivalent. For our purposes, it suffices to restrict our attention to pure state models.

Let \mathcal{Q} and \mathcal{Q}' be as in definition 4, with $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$ and $\rho'_\theta = |\psi'_\theta\rangle\langle\psi'_\theta|$ pure state on \mathcal{H} and respectively \mathcal{H}' . It has been shown [10] that the two models are equivalent if and only there exists a choice of phases such that the inner products coincide

$$\langle \psi_{\theta_1} | \psi_{\theta_2} \rangle = \langle \psi'_{\theta_1} | \psi'_{\theta_2} \rangle,$$

for all $\theta_1, \theta_2 \in \Theta$. This suggests the following definition of convergence of pure state statistical models.

Definition 5. A sequence of quantum statistical models $\mathcal{Q}_n := \{|\psi_\theta(n)\rangle\langle\psi_\theta(n)| : \theta \in \Theta\}$ over spaces \mathcal{H}_n converges weakly to a model $\mathcal{Q} := \{|\psi_\theta\rangle\langle\psi_\theta| : \theta \in \Theta\}$ over the space \mathcal{H} , if there exists a choice of phases for $|\psi_\theta(n)\rangle$ such that

$$\lim_{n \rightarrow \infty} \langle \psi_{\theta_1}(n) | \psi_{\theta_2}(n) \rangle = \langle \psi_{\theta_1} | \psi_{\theta_2} \rangle \quad (9)$$

for all $\theta_1, \theta_2 \in \Theta$.

Without entering into details we mention that this notion of convergence is closely related with that defined in [19], which in turn is a quantum analogue of the classical weak convergence of statistical models [33].

We introduce now a second notion of convergence of models which has a clearer operational interpretation and can be used to devise asymptotically optimal estimation procedures, and establish the asymptotic normality of estimators, in both the classical [33] and the quantum [28] contexts. For this, we will use definition 4 to build a distance between models, such that equivalent models have distance equal to zero, and small distance means that the models are ‘statistically close’ to each other but not necessarily equivalent.

Definition 6. With the notation of definition 4, we define the *Le Cam distance* between the models \mathcal{Q} and \mathcal{Q}' by $\Delta(\mathcal{Q}, \mathcal{Q}') = \max(\delta(\mathcal{Q}, \mathcal{Q}'), \delta(\mathcal{Q}', \mathcal{Q}))$ where

$$\delta(\mathcal{Q}, \mathcal{Q}') = \inf_T \sup_{\theta \in \Theta} \|T(\rho_\theta) - \rho'_\theta\|_1, \quad \delta(\mathcal{Q}', \mathcal{Q}) = \inf_S \sup_{\theta \in \Theta} \|S(\rho'_\theta) - \rho_\theta\|_1,$$

with T and S arbitrary quantum channels between the appropriated spaces.

A sequence of models $\mathcal{Q}_n := \{\rho_\theta(n) : \theta \in \Theta\}$ converges strongly to \mathcal{Q} if $\Delta(\mathcal{Q}, \mathcal{Q}_n) \rightarrow 0$ as $n \rightarrow \infty$.

It is easy to see that the strong convergence of a sequence \mathcal{Q}_n to \mathcal{Q} is equivalent to the existence of sequences (T_n) and (S_n) of channels such that

$$\lim_{n \rightarrow \infty} \sup_{\theta \in \Theta} \|T_n(\rho_\theta(n)) - \rho_\theta\|_1 = 0, \quad \text{and} \quad \lim_{n \rightarrow \infty} \sup_{\theta \in \Theta} \|S_n(\rho_\theta) - \rho_\theta(n)\|_1 = 0.$$

Indeed, assuming strong convergence, we find inductively for each $k \in \mathbb{N}$ an $n_k \geq n_{k-1}$ such that for each $n \geq n_k$ there exists a channel T_n^k with $\sup_{\theta \in \Theta} \|T_n^k(\rho_\theta(n)) - \rho_\theta\|_1 < 1/k$. Then the first of the above limits holds, if we define T_n for each $n \in \mathbb{N}$ by putting $T_n := T_n^k$ for that k for which $n \in [n_k, n_{k+1})$. The other relation follows similarly. The converse implication is clear.

The interpretation of the strong convergence of (quantum) statistical models is that asymptotically, the optimal risk associated to a statistical problem (e.g. estimation) for \mathcal{Q}_n converges to the optimal risk of the limit model; additionally, the optimal procedures can be related asymptotically via the channels T_n and S_n . A particular instance of strong convergence is *local asymptotic normality* of i.i.d. quantum models for mixed finite dimensional states [28], which shows that for large sample size the ensemble of identically prepared systems can be approximated via quantum channels by a multi-mode classical-quantum Gaussian state of unknown mean.

In this paper we establish a similar result for the correlated model given by the output state of a Markov chain. As a corollary we obtain the limiting (asymptotic) quantum Fisher information of the Markov output. This is done by proving the simpler weak convergence, which can be converted into strong convergence by applying Lemma 5 below. In preparation we prove a simpler lemma for finite number of parameters.

Lemma 4. *Let \mathcal{Q}_n and \mathcal{Q} be as in definition 5, and suppose that \mathcal{Q}_n converges weakly to \mathcal{Q} . Assume moreover that Θ is a finite set. Then \mathcal{Q}_n converges strongly to \mathcal{Q} , i.e. $\lim_{n \rightarrow \infty} \Delta(\mathcal{Q}_n, \mathcal{Q}) = 0$.*

Proof. By definition 5 we may assume that the phases of the vectors $|\psi_\theta(n)\rangle$ have been defined such that

$$\lim_{n \rightarrow \infty} \langle \psi_\theta(n) | \psi_\tau(n) \rangle = \langle \psi_\theta, \psi_\tau \rangle, \quad \forall \theta, \tau \in \Theta.$$

Let $\Theta = \{1, \dots, k\}$ and define the (positive) Gram matrix $G_{i,j}^{(n)} := \langle \psi_i(n) | \psi_j(n) \rangle$, and similarly $G_{i,j} := \langle \psi_i | \psi_j \rangle$. The convergence of the Gram matrices implies that for n large enough G_n and G have the same rank r .

Let $P^{(n)} := \sqrt{G^{(n)}}$, $P := \sqrt{G}$ and let $\{|e_1\rangle, \dots, |e_k\rangle\}$ be the standard orthonormal basis in \mathbb{C}^k . Then the vectors $|\tilde{\psi}_i(n)\rangle := \sum_{l=1}^k P_{i,l}^{(n)} |e_l\rangle \in \mathbb{C}^k$ have inner products

$$\langle \tilde{\psi}_i(n) | \tilde{\psi}_j(n) \rangle = ((P^{(n)})^* P^{(n)})_{i,j} = G_{i,j}^{(n)} = \langle \psi_i(n) | \psi_j(n) \rangle.$$

Similarly the vectors $|\tilde{\psi}_i\rangle := \sum_{l=1}^k P_{i,l} |e_l\rangle$ have inner products

$$\langle \tilde{\psi}_i | \tilde{\psi}_j \rangle = (P^* P)_{i,j} = G_{i,j} = \langle \psi_i | \psi_j \rangle.$$

Since two sets of vectors with the same Gram matrix are related by a single unitary matrix, we can map the models \mathcal{Q}_n and \mathcal{Q} into equivalent ones $\tilde{\mathcal{Q}}_n$ and respectively $\tilde{\mathcal{Q}}$, defined on the same representation space \mathbb{C}^k . By weak convergence we have that $G^{(n)}$ converges entrywise, and hence in norm to G . As the square root is a continuous function we get that $P^{(n)}$ converges to P and hence $|\tilde{\psi}_i(n)\rangle \rightarrow |\tilde{\psi}_i\rangle$ for all i . Therefore $\lim_{n \rightarrow \infty} \Delta(\tilde{\mathcal{Q}}_n, \tilde{\mathcal{Q}}) = 0$, and the same holds for \mathcal{Q}_n and \mathcal{Q} due to the above stated equivalence.

□

We will now strengthen this results, by assuming that we have more control over the weak convergence. The result will be used in upgrading the local asymptotic normality result in the next section, from weak to strong convergence.

Lemma 5. *Let Θ be a compact subset of \mathbb{R}^k , and let \mathcal{Q}_n and \mathcal{Q} be as in definition 5, with \mathcal{Q}_n converging weakly to \mathcal{Q} . Suppose, in addition, that the convergence in (9) is uniform for $(\theta_1, \theta_2) \in \Theta$. Then \mathcal{Q}_n converges strongly to \mathcal{Q} .*

The proof can be found in section 6.1.

5. LOCAL ASYMPTOTIC NORMALITY FOR QUANTUM MARKOV CHAINS

We now proceed to the second main task outlined in the introduction: characterising the statistical properties of the quantum output state of the Markov chain. We will assume that the Markov dynamics is unknown, or more precisely that $V = V_\theta$ depends smoothly on some unknown parameter $\theta \in \Theta \subset \mathbb{R}$. The case of multidimensional parameters can be investigated along the same lines, but will not be discussed here. A practical goal is to estimate θ by performing measurements on the output and constructing an estimator $\hat{\theta}$ based on the measurement results. The restriction to output measurements is justified by the fact that in many experiments the system cannot be measured directly, and is particularly natural in the context of quantum control engineering [34]. This problem was studied in [22] for a particular parametric family of discrete time quantum Markov chains, and the continuous time set-up was considered in [7]. The approach is based on tools of asymptotic statistics, which allows to distill the essential features of the problem such as asymptotic normality and Fisher information.

In this paper we focus on the properties of the quantum statistical model rather than studying particular measurement strategies, in the spirit of the classic work on quantum Cramér-Rao bound [3, 25, 24]. We first show that if the parameter θ is identifiable, the quantum statistical model $|\Psi_\theta(n)\rangle := |\Psi_{V_\theta}(n)\rangle$ (along with three other models describing the output) can be approximated by a quantum Gaussian shift model, and its limiting quantum Fisher information (per unit of time) can be computed analytically. On the other hand, for un-identifiable parameters such as in conjugation with a local unitary, we show that the corresponding Fisher information is zero.

The convergence to the Gaussian model is the extension to Markov dynamics of the Holstein-Primakov convergence for spin coherent states discussed in section 4.1. This involves a rescaling of the parameter with the statistical uncertainty such that $\theta := \theta_0 + u/\sqrt{n}$ with u the unknown local parameter. As already explained this does not amount to an additional assumption since for large n the ‘localisation’ can be achieved by means of an adaptive estimation procedure.

5.1. Quantum statistical models for estimation of Markov processes. We will analyse four (slightly) different definitions of the ‘output state’ associated to a primitive isometry V , and show that in asymptotics they lead to the same limit model and quantum Fisher information. Using the notations from subsection 2, we define the following states, where for the sake of clarity, the dependence on the initial system state and the isometry V is indicated explicitly:

- (a) The joint (pure) state of system and output and the memory after n iterations, for a given initial pure state $\varphi \in \mathcal{H}$ of the system

$$|\Psi_{V,\varphi}(n)\rangle := V(n)|\varphi\rangle;$$

- (b) The output state after n iterations, for a given initial pure state ϕ of the system

$$\rho_{V,\varphi}(n) := \text{tr}_{\mathcal{H}}[|\Psi_{V,\varphi}(n)\rangle\langle\Psi_{V,\varphi}(n)|]; \quad (10)$$

- (c) The output state corresponding to the stationary regime

$$\rho_V(n) := \text{tr}_{\mathcal{H}}[V(n)\rho_{ss}V(n)^*];$$

- (d) The (un-normalized) conditional state

$$|\psi_{V,\eta,\varphi}(n)\rangle := \sum_{\mathbf{i} \in I^{(n)}} \langle \eta | \mathbf{K}_{\mathbf{i}}^{(n)} | \varphi \rangle | \mathbf{i} \rangle \quad (11)$$

of the output after n iterations, on the condition that the initial state of the memory is $|\varphi\rangle$, and the measurement of the projection $|\eta\rangle\langle\eta|$ on the memory after the n iterations yields 1. In case $|\eta\rangle$ or $|\varphi\rangle$ is one of the basis vectors $|e_j\rangle$, we replace it by the index j so as to simplify the notation.

Here (c) is most natural from the operational point of view, since our basic physical assumption is that we only have access to the output in the stationary regime. The other states are, however, technically easier to handle; in particular, (d) and (a) have explicitly the form of a Matrix Product State.

Let us clarify the relations between the states (a) – (d); this will be useful later. Let $\rho_{ss} = \sum_{i=1}^D \Lambda_i |e_i\rangle\langle e_i|$ be the spectral decomposition of the stationary state ρ_{ss} . We can decompose the states (a) as

$$|\Psi_{V,\varphi}(n)\rangle = \sum_{j=1}^D |e_j\rangle \otimes |\psi_{V,j,\varphi}(n)\rangle, \quad (12)$$

and the state (b) is a mixture of states of type (d)

$$\rho_{V,\varphi}(n) = \sum_{j=1}^D |\psi_{V,j,\varphi}(n)\rangle \langle \psi_{V,j,\varphi}(n)|. \quad (13)$$

The stationary output state (c) is the mixture of pure states of type (d) corresponding to the system starting in one of the eigenstates $|e_i\rangle$, and ending in a state $|e_j\rangle$

$$\begin{aligned} \rho_V(n) &= \sum_{i=1}^D \Lambda_i \text{tr}_{\mathcal{H}}[|V(n)e_i\rangle \langle V(n)e_i|] \\ &= \sum_{i,j=1}^D \Lambda_i |\psi_{V,j,i}(n)\rangle \langle \psi_{V,j,i}(n)| \\ &= \sum_{i,j=1}^D \Lambda_i \Lambda_j |\Lambda_j^{-\frac{1}{2}} \psi_{V,j,i}(n)\rangle \langle \Lambda_j^{-\frac{1}{2}} \psi_{V,j,i}(n)|. \end{aligned} \quad (14)$$

5.2. Main results. In this section we present the second main result of the paper and discuss the interpretation of the associated quantum Fisher information. The proof is discussed in section 6. For the notations we refer to section 2.2.

Let $\theta \mapsto V_\theta$, be a smooth family of isometries parametrized by the unknown parameter $\theta \in \mathbb{R}$, and let $K_{i,\theta}$ be the corresponding Kraus operators, cf. section 2.1. We write $\theta = \theta_0 + u/\sqrt{n}$, with local parameter u and assume that $V := V_{\theta_0}$ is a primitive isometry, with stationary state ρ_{ss} . Let Λ_i denote its eigenvalues. We denote by $K_i := K_{i,\theta_0}$ and by \dot{K}_i the derivative of $K_{i,\theta}$ with respect to θ , at θ_0 . To emphasise the dependence on the local parameter we denote by $|\Psi_{u,\varphi}(n)\rangle$ the output state corresponding to $V_{\theta_0+u/\sqrt{n}}$, and similarly for the other states introduced in the previous subsection: $\rho_{u,\varphi}(n)$, $\rho_u(n)$ and $|\psi_{u,\eta,\varphi}(n)\rangle$. All the results in this section are based on the following Theorem 3, together with the observation that the relevant scalar products can clearly be written in the form

$$\langle \Psi_{u,\varphi}(n) | \Psi_{v,\varphi}(n) \rangle = \langle \varphi | T_{u,v,n}^n(\mathbb{1}) \varphi \rangle, \quad (15)$$

$$\langle \psi_{u,\eta,\varphi}(n) | \psi_{v,\eta,\varphi}(n) \rangle = \text{tr}[T_{u,v,n}^n(|\eta\rangle \langle \eta|) |\varphi\rangle \langle \varphi|], \quad (16)$$

where

$$T_{u,v;n} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}), \quad T_{u,v;n}(X) := V_{\theta_0+u/\sqrt{n}}^*(X \otimes \mathbb{1}) V_{\theta_0+v/\sqrt{n}}.$$

In order to simplify calculations we will additionally assume the following ‘gauge condition’

$$\text{Im} \sum_{i=1}^k \text{tr}[\rho_{ss} \dot{K}_i^* K_i] = 0. \quad (17)$$

It is straightforward to verify that (17) can always be satisfied by replacing the isometries $V = V_\theta$ with $e^{ib\theta} V_\theta$, where b is a suitable constant. Note that this transformation does not affect the actual statistical models given by the states $\Psi_{u,\phi}(n)$, $\rho_{u,\phi}(n)$, $\rho_u(n)$, and $\psi_{u,\eta,\phi}(n)$. It also does not change the equivalence class of the isometry V_θ .

Theorem 3. *There exist $F, a \in \mathbb{R}$, such that for each $X \in \mathcal{B}(\mathcal{H})$ and $C > 0$ we have*

$$\lim_{n \rightarrow \infty} \sup_{|u|, |v| < C} T_{u,v;n}^n(X) = \text{tr}[\rho_{ss} X] e^{-F(u-v)^2/8 - ia(u^2-v^2)} \mathbb{1}. \quad (18)$$

The constant F can be written in terms of the Markov variance of the ‘generator’ $G^* := i\dot{V}V^*$ (cf. section 2.2)

$$\begin{aligned} F = 4(G^*, G^*)_V &= 4\text{Tr} \{ \rho_{ss} \mathcal{E} [GG^* + 2\text{Re}[G(\mathbb{1} \otimes \mathcal{R} \circ \mathcal{E}(G^*))]] \} \\ &= 4 \sum_{i=1}^k \left[\text{Tr} \left[\rho_{ss} \dot{K}_i^* \dot{K}_i \right] + 2\text{Tr} \left[\text{Im}(K_i \rho_{ss} \dot{K}_i^*) \cdot \mathcal{R} \left(\text{Im} \sum_k \dot{K}_i^* K_i \right) \right] \right]. \end{aligned} \quad (19)$$

Proof. see section 6.2. □

5.2.1. *LAN for the pure system and output state.* Let $\mathcal{Q}_{n,\varphi} := \{|\Psi_{u,\varphi}(n)\rangle : u \in \mathbb{R}\}$ be the quantum statistical model consisting of system and output state $|\Psi_{V,\varphi}(n)\rangle$ (cf. section 5.1), with $V = V_{\theta_0+u/\sqrt{n}}$, and initial state $|\varphi\rangle$. We recall that

$$\mathcal{G} := \left\{ |\sqrt{F/2}u\rangle \in L^2(\mathbb{R}) : u \in \mathbb{R} \right\}$$

is a one-dimensional (pure states) quantum Gaussian model consisting of coherent states with mean $(\sqrt{F/2}u, 0)$.

Theorem 4. *The following convergence results hold:*

1. $\mathcal{Q}_{n,\varphi}$ converges weakly to \mathcal{G} ; more precisely, there exists a real constant a such that

$$\lim_{n \rightarrow \infty} \langle \exp(-iau^2) \Psi_{u,\varphi}(n) | \exp(-iav^2) \Psi_{v,\varphi}(n) \rangle = \langle \sqrt{F/2}u | \sqrt{F/2}v \rangle = \exp(-F(u-v)^2/8) \quad (20)$$

where $\exp(-iau^2)$ is a physically irrelevant phase factor, and F is the asymptotic quantum Fisher information per time unit, given by (19).

2. Let $Q_{n,\varphi}^c, G^c$ be the restrictions of $\mathcal{Q}_{n,\varphi}$ and \mathcal{G} to parameters in the bounded interval $u \in [-c, c]$. Then $Q_{n,\varphi}^c$ converges strongly to G^c , cf. definition 6.

Proof. The weak convergence follows immediately from Theorem 3 and (15). The strong convergence follows from the same Theorem 3 which shows that the above convergence is uniform over u, v in a bounded interval, and Lemma 5 which upgrades weak to strong convergence. □

5.2.2. *LAN for the pure conditional output state.* Since the system plus output state $|\Psi_{u,\varphi}(n)\rangle$ is pure (for a pure system initial state), its statistical properties are easier to understand, for which reason we have concentrated our analysis on this model so far. However, if we restrict ourselves to output measurements, it is natural to consider the output state alone, and its Fisher information. Intuitively, since in the long time limit the system reaches stationarity, while the Fisher information of the output state grows linearly, one expects that all information is in the output and no loss of information (per time unit) is incurred by tracing out the system.

We consider the output states $|\psi_{u,\eta,\varphi}(n)\rangle$, corresponding to an initial state $|\varphi\rangle$, and a final state $|\eta\rangle$ after n iterations. Let $\mathcal{Q}_{n,\eta,\varphi} := \{|\psi_{u,\eta,\varphi}(n)\rangle / \sqrt{N_n(u, \eta, \varphi)} : u \in \mathbb{R}\}$ be the corresponding statistical model, where $N_n(u, \eta, \varphi) = \|\psi_{u,\eta,\varphi}(n)\|^2$ is the normalisation constant.

Theorem 5. *The following convergence results hold:*

1. $\mathcal{Q}_{n,\eta,\varphi}$ converges weakly to \mathcal{G} .
2. Let $Q_{n,\eta,\varphi}^c, G^c$ be the restrictions of $\mathcal{Q}_{n,\eta,\varphi}$ and \mathcal{G} to parameters in the bounded interval $u \in [-c, c]$. Then $Q_{n,\eta,\varphi}^c$ converges strongly to G^c , cf. definition 6.

Proof. It follows immediately from Theorem 3 and (16) that

$$\lim_{n \rightarrow \infty} \langle \psi_{u,\eta,\varphi}(n) | \psi_{v,\eta,\varphi}(n) \rangle = \langle \eta | \rho_{ss} \eta \rangle \langle \sqrt{F/2}u, \sqrt{F/2}v \rangle,$$

which also implies that $\lim_{n \rightarrow \infty} N_n(u, \eta, \varphi) = \langle \eta | \rho_{ss} \eta \rangle$. This proves the weak convergence, and the strong convergence again follows from Theorem 3 and Lemma 5. □

5.2.3. *LAN for mixed output state with known initial state $|\eta\rangle$.* In a more realistic setup, we do not have access to the system after n iterations. Assuming that we know the initial state $|\varphi\rangle$, the output is described by the mixed state statistical model $\mathcal{O}_{n,\varphi} := \{\rho_{u,\varphi}(n) : u \in \mathbb{R}\}$, and its restriction to local parameters $u \in [-c, c]$ is denoted $\mathcal{O}_{n,\varphi}^c$. Since our definition of weak convergence applies only to pure states models, we formulate local asymptotic normality as strong converge to the Gaussian model.

Theorem 6. *The sequence of statistical models $\mathcal{O}_{n,\varphi}^c$ converges strongly to G^c .*

Proof. We use the decomposition (13) for $\rho_{u,\varphi}(n)$, with the eigenbasis of the stationary state ρ_{ss} of V :

$$\rho_{u,\varphi}(n) = \sum_{j=1}^D |\psi_{u,j,\varphi}(n)\rangle \langle \psi_{u,j,\varphi}(n)|.$$

Let

$$\tilde{\mathcal{O}}_{n,\varphi} := \{|\psi_{u,j,\varphi}(n)\rangle / \sqrt{N_n(u,j,\varphi)} : (u,j) \in \mathbb{R} \times \{1, \dots, D\}\}$$

be the ‘extended’ model consisting of all pure components $|\psi_{u,j,\varphi}(n)\rangle$, where the index j is seen as a discrete parameter. Again by using Theorem 3 and (16), we see that the extended model converges weakly to the extended Gaussian model

$$\tilde{G} := \{|\sqrt{F/2}u\rangle \otimes |j\rangle : (u,j) \in \mathbb{R} \times \{1, \dots, D\}\}$$

where $|\sqrt{F/2}u\rangle \otimes |j\rangle$ denotes a copy of the coherent state $|\sqrt{F/2}u\rangle$ in the tensor product $L^2(\mathbb{R}) \otimes \mathbb{C}^D$. Moreover since the above convergence is uniform, we can again apply Lemma 5 to upgrade the weak convergence to strong convergence. This means that for each $c > 0$ there exist quantum channels \tilde{T}_n such that

$$\lim_{n \rightarrow \infty} \sup_{|u| \leq c} \max_{j=1, \dots, D} \|N_n(u,j,\varphi)^{-1} \tilde{T}_n(|\psi_{u,j,\varphi}(n)\rangle \langle \psi_{u,j,\varphi}(n)|) - |\sqrt{F/2}u\rangle \langle \sqrt{F/2}u| \otimes |j\rangle \langle j|\|_1 = 0. \quad (21)$$

Let P be the channel taking partial trace over the label space \mathbb{C}^D . Then

$$P \left(|\sqrt{F/2}u\rangle \langle \sqrt{F/2}u| \otimes \sum_{j=1}^D \Lambda_j |j\rangle \langle j| \right) = |\sqrt{F/2}u\rangle \langle \sqrt{F/2}u|,$$

so using the channel contractivity for P we get

$$\|T_n(\rho_{u,\varphi}(n)) - |\sqrt{F/2}u\rangle \langle \sqrt{F/2}u|\|_1 \leq \sum_{j=1}^D \left\| \tilde{T}_n(|\psi_{u,j,\varphi}(n)\rangle \langle \psi_{u,j,\varphi}(n)|) - \Lambda_j |\sqrt{F/2}u\rangle \langle \sqrt{F/2}u| \otimes |j\rangle \langle j| \right\|_1,$$

where $T_n = P \circ \tilde{T}_n$. Since $\lim_{n \rightarrow \infty} N_n(u,j,\varphi) = \Lambda_j$, it follows from (21) that

$$\lim_{n \rightarrow \infty} \sup_{|u| \leq c} \|T_n(\rho_{u,\varphi}(n)) - |\sqrt{F/2}u\rangle \langle \sqrt{F/2}u|\|_1 = 0.$$

The existence of channels S_n such that

$$\lim_{n \rightarrow \infty} \sup_{|u| \leq c} \|\rho_{u,\varphi}(n) - S_n(|\sqrt{F/2}u\rangle \langle \sqrt{F/2}u|)\|_1 = 0.$$

is proved in a similar fashion. \square

5.2.4. LAN for the stationary output state. We consider now the stationary output model $\mathcal{SO}_n := \{\rho_u(n) : u \in \mathbb{R}\}$. For these states the decomposition (14) holds. We denote by \mathcal{SO}_n^c the restriction of \mathcal{SO}_n to parameters $u \in [-c, c]$, and we will show that \mathcal{SO}_n^c converges strongly to the (restricted) Gaussian model \mathcal{G}^c . This means that asymptotically, the stationary state carries the same amount of (Fisher) information as the system-output state which is a priori more informative since the observer has access to both system and output. The two models are asymptotically equivalent with a pure state Gaussian model.

Theorem 7. *For any $c > 0$, the stationary output state statistical model \mathcal{SO}_n^c converges strongly to \mathcal{G}^c .*

Proof. We use the decomposition (14) of $\rho_u(n)$:

$$\rho_u(n) = \sum_{j=1}^D \Lambda_j |\psi_{u,j,i}(n)\rangle \langle \psi_{u,j,i}(n)|.$$

Let

$$\widetilde{\mathcal{SO}}_n := \{|\psi_{j,i,u}(n)\rangle / \sqrt{N_n(u,j,i)} : (u,j,i) \in \mathbb{R} \times \{1, \dots, D\}^2\}$$

be the ‘extended’ model consisting of all pure components $|\psi_{j,i,u}(n)\rangle$, where the indices j, i are seen as discrete parameters. Again we can use Theorem 3 and (16) to conclude that this extended model converges weakly to the extended Gaussian model

$$\tilde{G} := \{|\sqrt{F/2}u\rangle \otimes |j,i\rangle : (u,j,i) \in \mathbb{R} \times \{1, \dots, D\}^2\},$$

where $|\sqrt{F/2}u\rangle \otimes |j,i\rangle$ denotes a copy of the coherent state $|\sqrt{F/2}u\rangle$ in the tensor product $L^2(\mathbb{R}) \otimes \mathbb{C}^{D^2}$. Moreover since the above convergence is uniform, we can apply Lemma 5 to upgrade the weak convergence to strong convergence. This means that there exist quantum channels \tilde{T}_n such that

$$\lim_{n \rightarrow \infty} \sup_{|u| \leq c} \max_{j,i=1, \dots, D} \|N_n(u,j,i)^{-1} \tilde{T}_n(|\psi_{j,i,u}(n)\rangle \langle \psi_{j,i,u}(n)|) - |\sqrt{F/2}u\rangle \langle \sqrt{F/2}u| \otimes |j,i\rangle \langle j,i|\|_1 = 0.$$

Let P be the channel taking partial trace over the label space \mathbb{C}^{D^2} , and take $T_n = P \circ \tilde{T}_n$. Since

$$P \left(|\sqrt{F/2}u\rangle \langle \sqrt{F/2}u| \otimes \sum_{j,i=1}^D \Lambda_j \Lambda_i |j,i\rangle \langle j,i| \right) = |\sqrt{F/2}u\rangle \langle \sqrt{F/2}u|,$$

we conclude, as in the proof of Theorem 6, that

$$\lim_{n \rightarrow \infty} \sup_{|u| \leq c} \|T_n(\rho_u(n)) - |\sqrt{F/2}u\rangle\langle\sqrt{F/2}u|\|_1 = 0.$$

Again the other limit needed for the strong convergence is proved in a similar fashion. \square

5.3. Unidentifiable parameters have zero Fisher information. If the Fisher information F at θ_0 is strictly positive, the above theorems 4, 5, 6, and 7 imply that θ can be estimated with standard mean square error rate $1/n$ in the neighbourhood of θ_0 , using any of the models considered.

On the other hand, if the isometries V_θ in the family are all equivalent in the sense of Definition 2, the associated *output* quantum Fisher information for such parameters must be strictly zero for outputs at arbitrary time (i.e. length n). However, for a given initial state φ of the system, the *joint* system plus output state *can* depend on the ‘hidden’ unitary conjugation and therefore may carry some information about it. Nevertheless, this information is not extensive in time, since the dynamics reaches stationarity in a finite time, and hence we would not expect it to contribute to the limiting Fisher information. Indeed, according to the results of the preceding section, the output model has the same limiting Fisher information per unit of time as the joint system plus output state.

The next corollary provides a consistency check, showing that any family of equivalent isometries has indeed zero limiting Fisher information for all of the statistical models considered above. The proof of this corollary can be found in Section 6.3.

Corollary 1. *Let V_θ , $\theta \in \Theta \subset \mathbb{R}$ be an analytic family of isometries, all belonging to the same equivalence class (c.f. Definition 2). Then the quantum Fisher information $F = F(\theta_0)$ associated to V_{θ_0} is equal to zero for all θ_0 .*

6. PROOFS

In this section we collect the proofs of various results of the paper.

6.1. Proof of Lemma 5. Denote $f_n(\theta_1, \theta_2) = |\langle\psi_{\theta_1}(n)|\psi_{\theta_2}(n)\rangle|^2$, and $f(\theta_1, \theta_2) = |\langle\psi_{\theta_1}|\psi_{\theta_2}\rangle|^2$ for all $\theta_1, \theta_2 \in \Theta$. Let us also denote the states $\rho_\theta(n) := |\psi_\theta(n)\rangle\langle\psi_\theta(n)|$ and $\rho_\theta := |\psi_\theta\rangle\langle\psi_\theta|$.

We prove that $\lim_{n \rightarrow \infty} \delta(\mathcal{Q}_n, \mathcal{Q}) = 0$. The other limit $\lim_{n \rightarrow \infty} \delta(\mathcal{Q}, \mathcal{Q}_n) = 0$ can be proved in a similar manner.

Let $\epsilon > 0$. According to the definition of strong convergence, we have to show that there exists an n_0 such that for each $n \geq n_0$ we have

$$\sup_{\theta \in \Theta} \|T_n(\rho_\theta(n)) - \rho_\theta\|_1 < \epsilon$$

for some channels T_n (which may of course also depend on ϵ).

Now by the assumption, the sequence (f_n) of functions converges to f uniformly on $\Theta \times \Theta$. Consequently, f is continuous and hence uniformly continuous due to compactness of Θ , so a simple $\epsilon/3$ -argument shows that $(f_n)_{n \geq n_1}$ is a *uniformly equicontinuous* family for some $n_1 \in \mathbb{N}$. This means that there exists a $\delta > 0$ such that

$$\sup_{n \geq n_1} |f_n(\theta_1, \theta_2) - f_n(\theta'_1, \theta'_2)| < \epsilon^2/36$$

whenever $|\theta_1 - \theta'_1| < \delta$ and $|\theta_2 - \theta'_2| < \delta$. Since $f_n(\theta, \theta) = 1$, we get, in particular, that

$$\|\rho_\theta(n) - \rho_\tau(n)\|_1 = 2\sqrt{1 - |\langle\psi_\theta(n)|\psi_\tau(n)\rangle|^2} = 2\sqrt{|f_n(\theta, \tau) - f_n(\theta, \theta)|} < \epsilon/3 \quad (22)$$

whenever $n \geq n_1$ and $|\theta - \tau| < \delta$. Moreover, (f_n) converges to f pointwise, so

$$\|\rho_\theta - \rho_\tau\|_1 = 2\sqrt{|f(\theta, \tau) - f(\theta, \theta)|} = \lim_{n \rightarrow \infty} 2\sqrt{|f_n(\theta, \tau) - f_n(\theta, \theta)|} \leq \epsilon/3 \quad (23)$$

whenever $|\theta - \tau| < \delta$.

Since Θ is compact, there exists a finite set $I_\delta \subset \Theta$ such that

$$\Theta \subset \bigcup_{\tau \in I_\delta} B_\delta(\tau), \quad (24)$$

where $B_\delta(\tau)$ denotes the open ball of radius δ centered at τ . The restrictions of \mathcal{Q}_n and \mathcal{Q} to I_δ are now finite state models. In this case weak convergence is equivalent with strong convergence as shown in Lemma 4, so there exist channels T_n^δ such that

$$\lim_{n \rightarrow \infty} \sup_{\tau \in I_\delta} \|T_n^\delta \rho_\tau(n) - \rho_\tau\|_1 = 0.$$

Consequently, we can find an $n_0 \geq n_1$ such that

$$\|T_n^\delta \rho_\tau(n) - \rho_\tau\|_1 < \epsilon/3 \quad \text{for all } n \geq n_0, \tau \in I_\delta. \quad (25)$$

Now for arbitrary $\theta \in \Theta$, there exists a $\tau \in I_\delta$ such that $|\theta - \tau| < \delta$ due to (24). With such a τ , the inequalities (22), (23), and (25) hold if $n \geq n_0$. Hence, using the fact that every channel has norm one, we get

$$\|T_n^\delta \rho_\theta(n) - \rho_\theta\|_1 \leq \|T_n^\delta(\rho_\theta(n) - \rho_\tau(n))\|_1 + \|T_n^\delta \rho_\tau(n) - \rho_\tau\|_1 + \|\rho_\tau - \rho_\theta\|_1 < \epsilon,$$

whenever $n \geq n_0$. This completes the proof. \square

6.2. Proof of Theorem 3. By analyticity, we can expand the Kraus operators using the Taylor series at θ_0 :

$$K_{i,\theta} = K_i + (\theta - \theta_0)\dot{K}_i + \frac{1}{2}(\theta - \theta_0)^2\ddot{K}_i + O((\theta - \theta_0)^3).$$

Hence, for a fixed $C > 0$, we have

$$K_{i,\theta_0+u/\sqrt{n}} = K_i + \frac{1}{\sqrt{n}}u\dot{K}_i + \frac{1}{2}\frac{1}{n}u^2\ddot{K}_i + O(n^{-\frac{3}{2}}).$$

uniformly for $|u| < C$, in the sense that the error term can be bounded above by a constant times $n^{-\frac{3}{2}}$. Using the fact that

$$T_{u,v;n}(X) = \sum_i K_{i,\theta_0+u/\sqrt{n}}^* X K_{i,\theta_0+v/\sqrt{n}}$$

it then follows that, again uniformly for $|u| < C$, $|v| < C$, we have

$$T_{u,v;n} = T + \frac{1}{\sqrt{n}}T_{1,u,v} + \frac{1}{n}T_{2,u,v} + O(n^{-3/2}), \quad (26)$$

where T is the channel associated to V , and

$$\begin{aligned} T_{1,u,v}(X) &= \sum_{i=1}^k \left\{ u\dot{K}_i^* X K_i + vK_i^* X \dot{K}_i \right\} \\ &= u\dot{V}^*(X \otimes \mathbb{1})V + vV^*(X \otimes \mathbb{1})\dot{V} \end{aligned} \quad (27)$$

$$\begin{aligned} T_{2,u,v}(X) &= \sum_{i=1}^k \left\{ \frac{u^2}{2}\ddot{K}_i^* X K_i + \frac{v^2}{2}K_i^* X \ddot{K}_i + uv\dot{K}_i^* X \dot{K}_i \right\} \\ &= \frac{u^2}{2}\ddot{V}^*(X \otimes \mathbb{1})V + \frac{v^2}{2}V^*(X \otimes \mathbb{1})\ddot{V} + uv\dot{V}^*(X \otimes \mathbb{1})\dot{V}. \end{aligned} \quad (28)$$

Using the fact that $\sum_i \dot{K}_i^* K_i = -\sum_i K_i^* \dot{K}_i$ by normalisation, we see that (17) implies

$$(T_{1,u,v}(\mathbb{1}), \mathbb{1})_{\theta_0} = i(u - v) \sum_{i=1}^k \text{Im}(\dot{K}_i^* K_i, \mathbb{1})_{\theta_0} = 0, \quad (29)$$

where $(A, B)_{\theta_0} := \text{Tr}(\rho_{ss} A^* B)$ defines an inner product on $\mathcal{B}(\mathcal{H})$ because $(A, A)_{\theta_0} = 0$ implies $\text{tr}[\rho_{ss} A^* A] = 0$ due to ρ_{ss} being full rank by the primitivity of T . Now (29) means that $T_{1,u,v}(\mathbb{1})$ belongs to $\mathcal{D}_1 := \{X : \text{Tr}(\rho_{ss} X) = 0\} = (\text{Id} - T)(\mathcal{B}(\mathcal{H}))$, which is the orthogonal complement of $\mathcal{D}_0 := \mathbb{C}\mathbb{1} = \ker(\text{Id} - T)$ with respect to the inner product $(\cdot, \cdot)_{\theta_0}$. Hence, the operator $\mathcal{R}(T_{1,u,v}(\mathbb{1}))$ is well defined, where \mathcal{R} is the inverse of the restriction of $(\text{Id} - T)$ to the invariant linear space \mathcal{D}_1 where that operator is bijective. This condition, together with the primitivity assumption for T , and the fact that each $T_{u,v;n}$ is a contraction in the operator norm, makes up the hypothesis of a second order perturbation theorem (Theorem 2) in [22]. Consequently, $\lim_{n \rightarrow \infty} \|T_{u,v;n}^n(\mathbb{1}) - \exp(\lambda(u, v))\mathbb{1}\| = 0$ holds for each u, v , where $\lambda(u, v)$ is given by

$$\lambda(u, v) = (\mathbb{1}, T_{2,u,v}(\mathbb{1}))_{\theta_0} + (\mathbb{1}, T_{1,u,v} \circ \mathcal{R} \circ T_{1,u,v}(\mathbb{1}))_{\theta_0}. \quad (30)$$

Moreover, since the perturbative expansion (26) (in powers of $n^{-\frac{1}{2}}$) is uniform for $|u|, |v| < C$ for any given $C > 0$, an inspection of the proof of the above mentioned theorem readily shows that, in fact,

$$\lim_{n \rightarrow \infty} \sup_{|u|, |v| \leq C} \|T_{u,v;n}^n(\mathbb{1}) - \exp(\lambda(u, v))\mathbb{1}\| = 0. \quad (31)$$

From (26) we also directly get the bound

$$\sup_{|u|, |v| \leq C} \|T_{u,v;n} - T\| \leq \sup_{|u|, |v| \leq C} \|T_{1,u,v}\| n^{-\frac{1}{2}} \leq \text{const.} n^{-\frac{1}{2}}. \quad (32)$$

As before we now write the spectral decomposition of ρ_{ss} as

$$\rho_{ss} = \sum_{i=1}^{\dim \mathcal{H}} \Lambda_i |e_i\rangle\langle e_i|$$

and we proceed to show the following statement

$$\lim_{n \rightarrow \infty} \sup_{|u|, |v| \leq C} \|T_{u,v;n}^n(|e_j\rangle\langle e_i|) - e^{\lambda(u,v)} \Lambda_i \delta_{i,j} \mathbb{1}\| = 0. \quad (33)$$

Let $k \in \mathbb{N}$, $k < n$. Since $T_{u,v;n}$ is a *contraction* in the operator norm, we can estimate

$$\begin{aligned} \|T_{u,v;n}^n(|e_j\rangle\langle e_i|) - e^{\lambda(u,v)} \Lambda_i \delta_{i,j} \mathbb{1}\| &\leq \|T_{u,v;n}^{n-k} T^k(|e_j\rangle\langle e_i|) - e^{\lambda(u,v)} \Lambda_i \delta_{i,j} \mathbb{1}\| \\ &+ \|T_{u,v;n}^k(|e_j\rangle\langle e_i|) - T^k(|e_j\rangle\langle e_i|)\| \\ &\leq \Lambda_i \delta_{i,j} \|T_{u,v;n}^{n-k}(\mathbb{1}) - e^{\lambda(u,v)} \mathbb{1}\| \\ &+ \|T^k(|e_j\rangle\langle e_i|) - \delta_{i,j} \Lambda_i \mathbb{1}\| \\ &+ \|T_{u,v;n}^k(|e_j\rangle\langle e_i|) - T^k(|e_j\rangle\langle e_i|)\| \\ &\leq \Lambda_i \delta_{i,j} (\|\mathbb{1} - T_{u,v;n}^k(\mathbb{1})\| + \|T_{u,v;n}^n(\mathbb{1}) - e^{\lambda(u,v)} \mathbb{1}\|) \\ &+ \|T^k(|e_j\rangle\langle e_i|) - \delta_{i,j} \Lambda_i \mathbb{1}\| \\ &+ \|T_{u,v;n}^k(|e_j\rangle\langle e_i|) - T^k(|e_j\rangle\langle e_i|)\|. \end{aligned}$$

For an arbitrary $\delta > 0$, we first choose k to be large enough so that $\|T^k(|e_j\rangle\langle e_i|) - \delta_{i,j} \Lambda_i \mathbb{1}\| < \delta/2$. This is possible because T_0 is primitive. With this fixed k , we have

$$\begin{aligned} \|\mathbb{1} - T_{u,v;n}^k(\mathbb{1})\| &= \|T^k(\mathbb{1}) - T_{u,v;n}^k(\mathbb{1})\| \leq k \|T - T_{u,v;n}\|, \\ \|T_{u,v;n}^k(|e_j\rangle\langle e_i|) - T^k(|e_j\rangle\langle e_i|)\| &\leq k \|T - T_{u,v;n}\|. \end{aligned}$$

Hence, using (31), and the bound (32), we obtain

$$\limsup_{n \rightarrow \infty} \sup_{|u|, |v| \leq C} \|T_{u,v;n}^n(|e_j\rangle\langle e_i|) - e^{\lambda(u,v)} \Lambda_i \delta_{i,j} \mathbb{1}\| < \delta,$$

so (33) holds.

It remains to show that

$$\lambda(u, v) = -\frac{F(u-v)^2}{8} - ia(u^2 - v^2)$$

with F given by (19). For this we compute now the two terms in (30). Firstly

$$\begin{aligned} (\mathbb{1}, T_{2,u,v}(\mathbb{1}))_{\theta_0} &= \frac{u^2 + v^2}{2} \sum_{i=1}^k \operatorname{Re}(\ddot{K}_i, K_i)_{\theta_0} + i \frac{u^2 - v^2}{2} \operatorname{Im}(\ddot{K}_i, K_i)_{\theta_0} \\ &+ uv \sum_{i=1}^k (\dot{K}_i, \dot{K}_i)_{\theta_0}. \end{aligned}$$

Note that $\sum_{i=1}^k K_{i;\theta}^* K_{i;\theta} = \mathbb{1}$ implies $\operatorname{Re}(\ddot{K}_i, K_i)_{\theta_0} = -(\dot{K}_i, \dot{K}_i)_{\theta_0}$ such that

$$(\mathbb{1}, T_{2,u,v}(\mathbb{1}))_{\theta_0} = -\frac{(u-v)^2}{2} \sum_{i=1}^k (\dot{K}_i, \dot{K}_i)_{\theta_0} + i \frac{u^2 - v^2}{2} \sum_{i=1}^k \operatorname{Im}(\ddot{K}_i, K_i)_{\theta_0}.$$

We pass now to the second term in (30). By differentiating $\sum_{i=1}^k K_{i;\theta}^* K_{i;\theta} = \mathbb{1}$ we get $\sum_{i=1}^k (\dot{K}_i^* K_i + K_i^* \dot{K}_i) = 0$ so that

$$T_{1,u,v}(\mathbb{1}) = (u-v) \sum_{i=1}^k \dot{K}_i^* K_i = i(u-v)(-i\dot{V}^* V). \quad (34)$$

This implies

$$(\mathbb{1}, T_{1,u,v} \circ \mathcal{R} \circ T_{1,u,v}(\mathbb{1}))_{\theta_0} = i(u-v) \operatorname{Tr} [T_{1,u,v}^*(\rho_{ss}) S],$$

where $S := \mathcal{R}(-i\dot{V}^* V)$. From (27) we get

$$T_{1,u,v}^*(\rho_{ss}) = (u+v) \operatorname{Re}(K_i \rho_{ss} \dot{K}_i^*) + i(u-v) \operatorname{Im}(K_i \rho_{ss} \dot{K}_i^*)$$

which we insert in the previous formula to obtain

$$\begin{aligned} (\mathbb{1}, T_{1,u,v} \circ \mathcal{R} \circ T_{1,u,v}(\mathbb{1}))_{\theta_0} = & - (u-v)^2 \sum_{i=1}^k \text{Tr}(\text{Im}(K_i \rho_{ss} \dot{K}_i^*) S) \\ & + i(u^2 - v^2) \sum_{i=1}^k \text{Tr}(\text{Re}(K_i \rho_{ss} \dot{K}_i^*) S) \end{aligned}$$

Finally, putting together the two terms we obtain

$$\begin{aligned} \lambda(u, v) = & - \frac{(u-v)^2}{2} \sum_{i=1}^k \left[(\dot{K}_i, \dot{K}_i)_{\theta_0} + 2 \text{Tr}(\text{Im}(K_i \rho_{ss} \dot{K}_i^*) S) \right] \\ & + i \frac{u^2 - v^2}{2} \sum_{i=1}^k \left[\text{Im}(\dot{K}_i, K_i)_{\theta_0} + 2 \sum_{i=1}^k \text{Tr}(\text{Re}(K_i \rho_{ss} \dot{K}_i^*) S) \right] \\ = & - \frac{(u-v)^2 F}{8} - ia(u^2 - v^2), \end{aligned}$$

where a is an unimportant constant, and F is the quantum Fisher information given by

$$F = 4 \sum_{i=1}^k \left[(\text{Tr}[\rho_{ss} \dot{K}_i^* \dot{K}_i] + 2 \text{Tr}(\text{Im}(K_i \rho_{ss} \dot{K}_i^*) S)) \right]. \quad (35)$$

We will now show that F can be expressed as a four times the variance of the (Markov) ‘generator’ $G^* := i\dot{V}V^*$, in analogy to the case of a unitary family of pure states. We have

$$V^* G V = -i V^* V \dot{V}^* V = -i \dot{V}^* V = i V^* \dot{V} = V^* G^* V,$$

because $\dot{V}^* V + V^* \dot{V} = 0$. In particular, $V^* G V$ is selfadjoint. We can now write $S = \mathcal{R}(V^* G^* V)$, which is also selfadjoint because $\text{Id} - T$, and hence also \mathcal{R} , commutes with the map $A \mapsto A^*$. Now

$$\begin{aligned} \sum_{i=1}^k \text{Tr}(\text{Im}(K_i \rho_{ss} \dot{K}_i^*) S) &= \sum_{i=1}^k \text{Im} \text{Tr}(K_i \rho_{ss} \dot{K}_i^* S) \\ &= \text{Im} \text{Tr}(\rho_{ss} \dot{V}^* (\mathbb{1} \otimes S) V) \\ &= \text{Im} \text{Tr}(\rho_{ss} V^* i G (\mathbb{1} \otimes S) V) \\ &= \text{Re} \text{Tr}[\rho_{ss} \mathcal{E}(G(\mathbb{1} \otimes \mathcal{R} \circ \mathcal{E}(G^*)))]. \end{aligned}$$

On the other hand

$$\sum_{i=1}^k \text{Tr}[\rho_{ss} \dot{K}_i^* \dot{K}_i] = \text{Tr}(\rho_{ss} V^* V \dot{V}^* \dot{V} V^* V) = \text{Tr}(\rho_{ss} \mathcal{E}(G G^*)).$$

Therefore

$$F = 4(G^*, G^*)_V = 4 \text{Tr} \{ \rho_{ss} \mathcal{E} [G G^* + 2 \text{Re}[G(\mathbb{1} \otimes \mathcal{R} \circ \mathcal{E}(G^*))]] \}.$$

6.3. Proof of Corollary 1.

Proof. Fix $\theta_0 \in \mathbb{R}$, and denote $V = V_{\theta_0}$, with also other notations as before. According to Theorem 2, for each $\theta \in \Theta$ there exists a unitary operator \tilde{U}_θ on \mathcal{H} , and a phase factor $e^{if(\theta)}$ such that

$$V_\theta = e^{if(\theta)} (\tilde{U}_\theta \otimes \mathbb{1}) V \tilde{U}_\theta^*. \quad (36)$$

Since the statistical models associated to V_θ are not affected by multiplication by a phase factor, we may assume without loss of generality that $f = 0$. Now we have $V = (\tilde{U}_{\theta_0} \otimes \mathbb{1}) V \tilde{U}_{\theta_0}^*$, which implies that $K_i \tilde{U}_{\theta_0} = \tilde{U}_{\theta_0} K_i$ for each Kraus operator K_i . Since V is a primitive isometry, it follows that for a large enough n , the Kraus operators $\mathbf{K}_1^{(n)}$ span the whole space $\mathcal{B}(\mathcal{H})$ (see e.g. [32]). Hence $X = \tilde{U}_{\theta_0} X \tilde{U}_{\theta_0}^*$ for all $X \in \mathcal{B}(\mathcal{H})$. From this it follows that $\tilde{U}_{\theta_0} = e^{ig(\theta_0)} \mathbb{1}$, where $e^{ig(\theta)} := \langle \mathbb{1} | \tilde{U}_\theta | \mathbb{1} \rangle$. Similarly, since $\theta \mapsto V_\theta$ is analytic, it follows that $\theta \mapsto \tilde{U}_\theta K_i \tilde{U}_\theta^*$ is analytic for all Kraus operators K_i , so again by primitivity, $\theta \mapsto \tilde{U}_\theta X \tilde{U}_\theta^*$ is analytic for all $X \in \mathcal{B}(\mathcal{H})$.

Hence, the unitary family $\theta \mapsto U_\theta := e^{-ig(\theta)} \tilde{U}_\theta$ is analytic, satisfies $U_{\theta_0} = \mathbb{1}$, and we have $V_\theta = (U_\theta \otimes \mathbb{1}) V U_\theta^*$.

Since U_θ is an analytic family, we may differentiate; we let \dot{U} denote the derivative of U_θ at $\theta = \theta_0$, and put

$$H := -i\dot{U}^*.$$

Since $U_\theta^* U_\theta = \mathbb{1}$, and $U_0 = \mathbb{1}$, we have $\dot{U}^* + \dot{U} = 0$, which implies $H = -i\dot{U}^* = i\dot{U} = H^*$, i.e. H is selfadjoint. Hence,

$$i\dot{V} = i(\dot{U} \otimes \mathbb{1})V + iV\dot{U}^* = (\mathbb{1} \otimes H)V - VH. \quad (37)$$

We first verify that $(T_{1,u,v}(\mathbb{1}), \mathbb{1})_{\theta_0} = 0$ so no phase adjustment is needed for the Kraus operators. Indeed, using equation (34) from the preceding proof, we get

$$\begin{aligned} (\mathbb{1}, T_{1,u,v}(\mathbb{1}))_{\theta_0} &= \text{tr}[\rho_{ss} T_{1,u,v}(\mathbb{1})] = i(u-v)\text{tr}[\rho_{ss}(-i\dot{V}^*V)] \\ &= i(u-v)\text{tr}[\rho_{ss}(V^*(\mathbb{1} \otimes H)V - HV^*V)] = i(u-v)\text{tr}[\rho_{ss}(T(H) - H)] = 0, \end{aligned}$$

because ρ_{ss} is the stationary state. In order to find the Fisher information we compute

$$\begin{aligned} G^* &= i\dot{V}V^* = (\mathbb{1} \otimes H)VV^* - VHV^*, \\ \mathcal{E}(G^*) &= V^*G^*V = V^*(\mathbb{1} \otimes H)V - VH = T(H) - H = -(\text{Id} - T)(H), \\ \mathcal{R} \circ \mathcal{E}(G^*) &= -H, \\ \mathcal{E}(G(\mathbb{1} \otimes \mathcal{R} \circ \mathcal{E}(G^*))) &= -\mathcal{E}(G(\mathbb{1} \otimes H)) = -V^*(VV^*(\mathbb{1} \otimes H) - VHV^*)(\mathbb{1} \otimes H)V \\ &= -T(H^2) + HT(H), \\ \mathcal{E}(2\text{Re}((G(\mathbb{1} \otimes \mathcal{R} \circ \mathcal{E}(G^*)))) &= 2\text{Re}\mathcal{E}(G(\mathbb{1} \otimes \mathcal{R} \circ \mathcal{E}(G^*))) = 2\text{Re}(-T(H^2) + HT(H)) \\ &= -2T(H^2) + HT(H) + T(H)H, \\ \mathcal{E}(GG^*) &= V^*(VV^*(\mathbb{1} \otimes H) - VHV^*)((\mathbb{1} \otimes H)VV^* - VHV^*)V \\ &= (V^*(\mathbb{1} \otimes H) - HV^*)((\mathbb{1} \otimes H)V - VH) \\ &= V^*(H^2 \otimes \mathbb{1})V - HV^*(\mathbb{1} \otimes H)V - V^*(\mathbb{1} \otimes H)VH + H^2 \\ &= T(H^2) - HT(H) - T(H)H + H^2. \end{aligned}$$

Hence,

$$\begin{aligned} F &= 4\text{tr}\{\rho_{ss}\mathcal{E}[GG^* + 2\text{Re}[G(\mathbb{1} \otimes \mathcal{R} \circ \mathcal{E}(G^*))]]\} \\ &= 4\text{tr}[\rho_{ss}(-T(H^2) + H^2)] = 0, \end{aligned}$$

because ρ_{ss} is the stationary state. □

7. CONCLUSIONS

We have shown that the dynamics of an ergodic quantum Markov chain can be identified up to a ‘change of coordinates’ transformation consisting of a phase multiplication and a unitary conjugation of the Kraus operators. This result can be seen as a quantum analogue of the classical result by Petrie [30] on hidden Markov chains. An interesting open problem is to find an explicit algorithm for computing the equivalence class from the output state.

We have further shown that the output state, which is a ‘purely generated’ finitely correlated state [15], satisfies the local asymptotic normality property, and we have provided an explicit expression of the limiting quantum Fisher information, which can be interpreted as a Markov variance of a certain generator, in analogy to the case of unitary rotation families of quantum state. This analysis can be pushed forward by analysing the statistical properties of certain measurements such as simple repeated measurements on the output [22]. In general such measurements do not achieve the quantum Fisher information and it remains an open question to characterise the optimal measurement procedure. This is related to another open problem, that of establishing a general quantum Central Limit for the fluctuation operators $\mathbb{F}_n(X)$ defined in section 2.2.

More generally, we conjecture that local asymptotic normality holds for Markov chains with mixed i.i.d. input states. On the other hand, Markov chains with several ergodic components may fail to satisfy local asymptotic normality, and even exhibit ‘Heisenberg scaling’, with relevance for quantum metrology applications.

Acknowledgment. This work was supported by the EPSRC project EP/J009776/1.

REFERENCES

- [1] B. Baumgartner and H. Narnhofer. Analysis of quantum semigroups with GKSLindblad generators: II. General. *J. Phys. A: Math. Theor.*, 41:395303, 2008.
- [2] T. Baumgratz, D. Gross, M. Cramer, and M.B. Plenio. Scalable reconstruction of density matrices. *Phys. Rev. Lett.*, 111:020401, 2013.
- [3] V. P. Belavkin. Generalized Heisenberg uncertainty relations, and efficient measurements in quantum systems. *Theor. Math. Phys.*, 26:213–222, 1976.
- [4] P.J. Bickel, Y. Ritov, and T. Ryden. Inference in hidden Markov models I: Local asymptotic normality in the stationary case. *Bernoulli*, 2:199–291, 1996.
- [5] P.J. Bickel, Y. Ritov, and T. Ryden. Asymptotic normality of the Maximum likelihood estimator for general hidden Markov models. *The Annals Statistics*, 26:1614–1635, 1998.
- [6] D. Burgarth, K. Maruyama, and F. Nori. Indirect quantum tomography of quadratic Hamiltonians. *New J. of Phys.*, 13:013019, 2011.
- [7] C. Catana, M. Guță, and L. Bouten. Local asymptotic normality and Fisher information for quantum Markov processes. in preparation.
- [8] C. Catana, M. Guță, and T. Kypraios. Maximum likelihood versus likelihood-free quantum system identification in the atom maser. arXiv:1311.4091, 2013.
- [9] C. Catana, M. van Horssen, and M. Guță. Asymptotic inference in system identification for the atom maser. *Phil. Trans. R. Soc. Lond. A*, 370:5308–5323, 2012.
- [10] A. Chefles, R. Josza, and A. Winter. On the existence of physical transformations between sets of quantum states. *Int. J. Quantum Inform.*, 02:11, 2004.
- [11] J. Cole, S. Schirmer, A. Greentree, C. Wellard, D. Oi, and L. Hollenberg. Identifying an experimental two-state Hamiltonian to arbitrary accuracy. *Phys. Rev. A*, 71:062312, 2005.
- [12] R. Douc and C. Matias. Asymptotics of the Maximum Likelihood Estimator for general Hidden Markov Models. *Bernoulli*, 7:381–420, 2001.
- [13] J. P. Dowling and G. J. Milburn. Quantum technology: the second quantum revolution. *Phil. Trans. R. Soc. Lond. A*, 361:1655–1674, 2003.
- [14] D.E. Evans and R. Hoegh-Krohn. Spectral properties of positive maps on c^* -algebras. *J. London Math. Soc.*, s2-17:345, 1978.
- [15] M. Fannes, B. Nachtergaele, and R. F. Werner. Finitely correlated states on quantum spin chains. *Commun. Math. Phys.*, 144:443, 1992.
- [16] A. Fujiwara. Quantum channel identification problem. *Phys. Rev. A*, 63:042304, 2001.
- [17] C. Gardiner and P. Zoller. *Quantum Noise*. Springer, 2004.
- [18] R. D. Gill and M. Guță. On Asymptotic Quantum Statistical Inference. *Institute of Mathematical Statistics Collections*, 9:105–127, 2012.
- [19] M. Guță and A. Jenčová. Local asymptotic normality in quantum statistics. *Commun. Math. Phys.*, 276:341–379, 2007.
- [20] M. Guță and J. Kahn. Local asymptotic normality for qubit states. *Phys. Rev. A*, 73:052108, 2006.
- [21] M. Guță and M. van Horssen. Large Deviations, Central Limit and dynamical phase transitions in the atom maser. arXiv:1206.4956v2, 2012.
- [22] M. Guță. Fisher information and asymptotic normality in system identification for quantum Markov chains. *Phys. Rev. A*, 83:062324, 2011.
- [23] P. Haroche and J.-M. Raimond. *Exploring the Quantum: Atoms, Cavities, and Photons*. Oxford University Press, 2006.
- [24] C. W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.
- [25] A. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland Publishing Company, Amsterdam-New York-Oxford, 1982.
- [26] T. Holstein and H. Primakoff. Field dependence of the intrinsic domain magnetization of a ferromagnet. *Phys. Rev.*, 58:1098–1113, 1940.
- [27] M. Howard, J. Twamley, C. Wittmann, T. Gaebel, F. Jelezko, and J. Wrachtrup. Quantum process tomography and Linblad estimation of a solid-state qubit. *New J. of Phys.*, 8:33, 2006.
- [28] J. Kahn and M. Guță. Local asymptotic normality for finite dimensional quantum systems. *Commun. Math. Phys.*, 289:597–652, 2009.
- [29] H. Mabuchi and N. Khaneja. Principles and applications of control in quantum systems. *Int. J. Robust Nonlinear Control*, 15:647–667, 2005.
- [30] T. Petrie. Probabilistic functions of finite state Markov chains. *Annals of Math. Statistics*, 40:97–115, 1969.
- [31] D. Petz and Jencova, A. Sufficiency in quantum statistical inference. *Commun. Math. Phys.*, 263:259 – 276, 2006.
- [32] M. Sanz, D. Pérez-García, M.M. Wolf, and J.I. Cirac. A Quantum version of Wielandt’s inequality. *IEEE Trans. Inform. Theory*, 56:4668–4673, 2010.
- [33] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.
- [34] H. M. Wiseman and G. J. Milburn. *Quantum Measurement and Control*. Cambridge University Press, 2009.

E-mail address: madalin.guta@nottingham.ac.uk

E-mail address: jukka.kiukas@nottingham.ac.uk

SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF NOTTINGHAM, UNIVERSITY PARK, NOTTINGHAM, NG7 2RD, UK